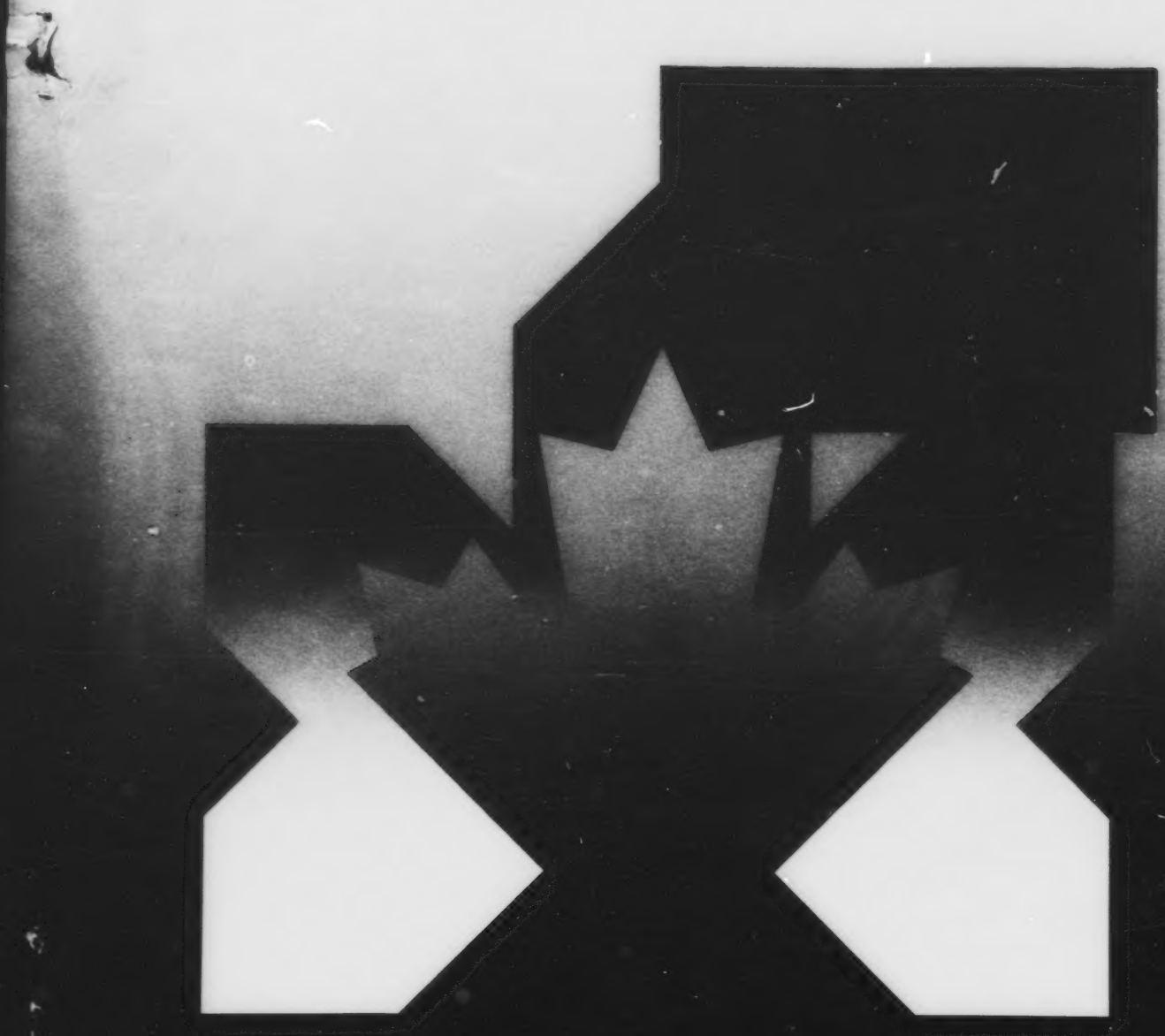
SCHL CMHC

Question habitation, comptez sur nous



Canada

93-01320/

LA PRÉVISION APPLIQUÉE AU SECTEUR DE L'HABITATION ET AU MARCHÉ HYPOTHÉCAIRE: EXAMEN MÉTHODOLOGIQUE: RAPPORT FINAL

SOO-BIN PARK ET N. HARVEY LITHWICK

UNIVERSITÉ CARLETON

MARS 1991

Le travail qui fait l'objet du présent rapport a été commandé par la Société canadienne d'hypothèques et de logement. Nous remercions Mark Burchinshaw et Leonard FitzPatrick pour leurs nombreux et utiles commentaires et suggestions. Nous remercions également Peter Torraville et Geoffrey Oliver pour leur assistance de recherche compétente. Peter Torraville a exécuté la plus grande partie du travail à l'ordinateur réalisé au cours du projet.

TABLE DES MATIÈRES

1. Introduction

2. Prévoir avec un modèle économétrique

2.1 Le modèle économétrique à équations simultanées

2.2 Prévoir avec un modèle économétrique

2.3 Le modèle du secteur du logement

2.3.1 Théorie économique du logement

2.3.2 Cadre du macromodèle du secteur du logement

2.3.3 Cadre du micromodèle du marché

2.3.4 Les données

3. Prévoir avec un modèle univarié

3.1 Un modèle ARIMA des mises en chantier

3.2 Prévoir avec un modèle ARIMA

4. Prévoir avec un modèle à fonction de transfert

4.1 Un modèle à fonction de transfert des mises en chantier

4.2 Prévoir avec un modèle à fonction de transfert

5. Prévoir avec un modèle autorégressif vectoriel

5.1 Le système VAR de Sims

5.2 Le système VAR de Hsiao

5.3 Le système VAR bayésien

5.4 Efficacité prévisionnelle des modèles VAR

6. Conclusions

Références

1. INTRODUCTION

Les marchés domiciliaire et hypothécaire se caractérisent par des fluctuations cycliques et par le déséquilibre. La volatilité du secteur du logement apparaît clairement à la lumière du taux de croissance annuel des mises en chantier. Au Canada il a oscillé de -26,4 à 26,0 p. 100 au cours de la période de 1961 à 1989, pendant que le taux de croissance du produit domestique brut réel variait de -3,65 à 7,14 p. 100. Il est également couramment admis que le marché national du logement dans une économie consiste en de nombreux marchés locaux segmentés qui peuvent à court terme être en déséquilibre à cause, entre autres facteurs, de l'immobilité du parc résidentiel. De telles caractéristiques expliquent partiellement pourquoi il est difficile de bien prévoir l'activité des marchés domiciliaire et hypothécaire au Canada.

L'objectif premier de la présente étude est de passer en revue les principales méthodes de prévision susceptibles d'être employées pour prévoir au niveau macroéconomique les marchés domiciliaire et hypothécaire canadien. Ces dernières années, la plus grande partie de l'activité prévisionnelle s'est déployée au niveau macroéconomique au moyen de modèles économétriques à équations simultanées. Mais la valeur à court terme des modèles macroéconomiques comme outil prévisionnel est devenue l'objet de scepticisme; voir, par exemple, Stekler (1968) et Sims (1980). Box et Jenkins (1970) ont mis en vogue en dehors des cercles économiques une classe de modèles à séries temporelles servant d'outil prévisionnel. Les modèles à séries temporelles sont ainsi devenus des concurrents de taille pour les modèles économétriques traditionnels.

Nous nous attachons dans la présente étude à évaluer les principales méthodes de prévision axées sur un modèle. On peut classifier les modèles en deux groupes : ceux fondés sur les méthodes économétriques et ceux s'appuyant sur les séries temporelles. Parmi les modèles du premier groupe figurent :

(1) le modèle régressif linéaire multiple, fondé sur la relation, selon la théorie économique, entre une variable dépendante (ou variable effet) et un ensemble de variables indépendantes (variables causes); et
(2) le modèle économétrique à équations simultanées (SEM), qui comporte des équations structurelles décrivant les rapports entre les variables endogènes et les variables prédéterminées tels que spécifiés par la théorie économique pertinente.

Le second groupe de modèles comprend les différents modèles suivants :

(1) le modèle autorégressif moyenne mobile intégré univarié (ARIMA), souvent appelé modèle Box-Jenkins, est utilisé

pour effectuer des prévisions portant sur une série temporelle unique à partir de son évolution dans le passé;

(2) le modèle à fonction de transfert établit une relation entre une série temporelle (de sortie), soumise à prédiction, et un ensemble de variables (d'entrée); ce modèle permet de prévoir l'évolution d'une série temporelle non seulement à partir de son propre passé, mais également à partir de celui d'autres variables liées; ce modèle ressemble à la régression linéaire multiple; et

(3) le modèle autorégressif vectoriel moyenne mobile intégré (VARIMA) est une extension du modèle à variable unique ARIMA, puisqu'il incorpore une série temporelle multiple avec interaction ou rétroaction réciproques; l'autorégression vectorielle (VAR) en représente un cas particulier.

L'approche économétrique pure fait appel à la théorie économique pour élaborer un modèle économétrique; c'est pourquoi on dit qu'elle est « à base théorique ». Les données sont utilisées à l'étape de l'estimation du modèle. En revanche, l'approche par séries temporelles est largement tributaire des données pour l'élaboration du modèle; elle est donc « à base de données ». Bien entendu, l'analyste de séries temporelles s'appuie également sur la théorie économique pour y déterminer quelles variables pourraient être modélisées conjointement dans les modèles à fonction de transfert ou VAR.

Qu'on se serve à des fins prévisionnelles d'un modèle économétrique ou bien d'un modèle à série temporelle, il faut dans les deux cas considérer les questions suivantes pour élaborer et exploiter le modèle :

(1) spécification des équations du modèle - en fonction des objectifs du modèle et, dans le cas d'un modèle économétrique, des considérations théoriques pertinentes;
(2) collecte des données nécessaires;
(3) estimation initiale des paramètres inconnus du modèle; et
(4) validation du modèle - vérification diagnostique, simulation et tests de suivi du modèle estimé.

Il faut aborder ces questions avec beaucoup de soin si l'on veut obtenir des prévisions fiables à partir du modèle estimé.

Dans le présent rapport, on examine le modèle économétrique standard de même que les trois modèles à série temporelle, le modèle régressif étant considéré comme un cas spécial de modèle économétrique. De ces quatre modèles, deux sont de type systémique et les deux autres sont à équation unique. Les

modèles SEM standard ou VAR sont de type systémique en ce sens qu'ils renferment deux équations ou plus. Le modèle économétrique contient autant d'équations qu'il y a de variables endogènes et fournit des prévisions portant sur toutes les variables endogènes à partir des valeurs futures des variables prédéterminées. Par contre, le modèle VAR traite toutes les variables du système comme des variables endogènes et fournit des prévisions conjointes pour toutes les variables du système à partir de leurs valeurs présentes et passées.

Le modèle univarié ARIMA et le modèle à fonction de transfert sont à équation unique. Mais on peut les utiliser pour modéliser les marchés domiciliaire et hypothécaire en élaborant un système dans lequel chaque équation ne renferme qu'une seule variable endogène. Mills et Stephenson (1987) ont fourni un exemple d'application fructueuse de cette approche à un système prévisionnel à série temporelle de la masse monétaire au Royaume-Uni.

Le plan du présent rapport est le suivant. Au Chapitre 2 on examine le modèle économétrique à équations simultanées en tant qu'outil prévisionnel et on fournit le cadre conceptuel du modèle économique du marché domiciliaire et hypothécaire. On y passe également en revue certains modèles macroéconomiques canadiens selon la manière dont y est envisagé le secteur du logement et on examine la disponibilité des données.

Les Chapitres 3, 4 et 5 sont consacrés aux modèles à séries temporelles. Au Chapitre 3 nous décrivons le modèle ARIMA et présentons les résultats de prévisions fondées sur un modèle ARIMA saisonnier multiplicatif des mises en chantier au Canada. Le Chapitre 4 traite du modèle à fonction de transfert; on y présente en outre un modèle à fonction de transfert de mises en chantier, avec les indices du prix des logements comme variable d'entrée. Le Chapitre 5 est consacré à trois types différents de modèles VAR : les modèles de type Sims, Hsiao et bayésien. On y compare également les prévisions sur les mises en chantier fournies par divers types de modèles VAR bivariés des mises en chantier et des consentements hypothécaires. On trouvera au Chapitre 6 des remarques de conclusion.

2. PRÉVOIR AVEC UN MODÈLE ÉCONOMÉTRIQUE

Depuis l'oeuvre de pionniers de Tinbergen (1939) et de Klein (1950) sur la modélisation macroéconomique, le modèle économétrique à équations simultanées (SEM) est devenu un instrument très important à des fins de prévision et d'analyse décisionnelle. On a assisté pendant les années 1960 et 1970 au développement rapide des modèles macroéconomiques favorisé par le développement des méthodes économétriques, d'énormes progrès dans le pouvoir de calcul et la disponibilité de données économiques de bonne qualité. Dans le présent chapitre on fait l'examen du SEM en tant qu'outil de prévision.

Nous décrivons brièvement dans la section 2.1 le modèle économétrique standard à équations simultanées. Dans la section 2.2 sont décrites les principales caractéristiques de l'approche prévisionnelle au moyen d'un modèle économétrique. Dans la section 2.3 nous fournissons le cadre conceptuel pour le secteur du logement d'une économie et passons brièvement en revue les secteurs domiciliaires de certains modèles macroéconomiques et les modèles domiciliaires nationaux. Dans la section 2.4 sont examinées la disponibilité et la qualité des données sur les variables relatives au marché domiciliaire et hypothécaire.

2.1 LE MODÈLE ÉCONOMÉTRIQUE À ÉQUATIONS SIMULTANÉES

En théorie économique on émet souvent l'hypothèse qu'un ensemble de variables endogènes sont conjointement interdépendantes et simultanément déterminées par un ensemble de variables prédéterminées. On peut se représenter les variables endogènes comme les variables d'intérêt, alors que les variables prédéterminées comprennent des valeurs retardées de variables endogènes et des variables exogènes qu'on suppose déterminées à l'extérieur du modèle.

On représente généralement un <u>modèle économétrique à équations simultanées</u> au moyen d'un système de G équations structurelles contenant G variables endogènes, Y_1, Y_2, ..., Y_G, et K variables prédéterminées, X_1, X_2, ..., X_K. En pratique, de nombreuses équations structurelles sont non linéaires parce que diverses fonctions telles que rapports et logarithmes de certaines des variables apparaissent dans le modèle. Une expression générale de la $i^{ème}$ équation structurelle est

$$f_i(y_t, x_t; \theta_i) = u_{ti}, \quad i = 1, \ldots, G \quad t = 1, \ldots T \tag{2.1}$$

où f_i est une fonction générale connue, θ_i est un vecteur des coefficients structurels dans la $i^{ème}$ équation, $y'_t = (Y_{t1}, \ldots, Y_{tG})$ et $x'_t = (X_{t1}, \ldots, X_{tK})$ sont des vecteurs G et K

d'observations sur les variables endogènes $Y_1, \ldots, Y_G$ et sur les variables prédéterminées $X_1, \ldots, X_K$, respectivement, et u_{ti} est la perturbation aléatoire dans la $i^{ème}$ équation, tous au moment t. Les θ sont appelés coefficients structurels. La théorie économique fournit une information a priori sur les coefficients θ, de manière à ce qu'ils soient « identifiés ».

Lorsque le SEM est dynamique et que ses variables prédéterminées incluent des valeurs retardées de variables endogènes, il spécifie comment les chemins temporels des variables endogènes sont engendrés par les chemins temporels des variables exogènes. Supposons que les G premières variables prédéterminées dans le modèle soient des variables endogènes retardées d'une période et que les K - G variables prédéterminées restantes soient exogènes, de sorte que nous puissions segmenter le vecteur des observations sur les variables prédéterminées au temps t comme suit : $x'_t = (y'_{t-1} \vdots z'_t)$. On peut alors écrire le SEM en $(2.1)^t$ de la manière suivante :

$$f_i(y_t, y_{t-1}, z_t; \theta_i) = u_{ti} \quad i = 1, \ldots, G \quad t = 1, \ldots, T \tag{2.2}$$

Un modèle économétrique tel que (2.1) ou (2.2) comporte une représentation explicite des relations de causalité ou de rétroaction présumées entre les variables du modèle. C'est par ce biais que la théorie économique joue un rôle de premier plan en modélisation économétrique. Dans l'approche économétrique de la modélisation et de la prévision, la principale préoccupation du prévisionniste consiste à formuler un modèle qui soit dérivé de la théorie économique. Mais, pour plusieurs raisons, l'approche « à base théorique » pure pourra se révéler peu pratique.

En effet, la théorie économique peut fournir des hypothèses concurrentes. Dans le cas d'hypothèses emboîtées, nous pourrons choisir une hypothèse particulière en nous appuyant sur des tests statistiques standard sur les paramètres pour lesquels nous disposons d'hypothèses. On a consenti ces dernières années des efforts de recherche considérables sur la question des choix entre hypothèses concurrentes non emboîtées dans les modèles de régression. Voir à ce sujet, par exemple, Pesaran (1974), Fisher et McAleer (1981) et McAleer et Pesaran (1986). Toutefois, le concepteur de modèle ne peut en pratique montrer avec confiance une préférence pour un modèle ou l'autre.

La théorie pourra suggérer que sont pertinentes des variables pour lesquelles aucune donnée n'est disponible. Ainsi, des variables telles que le revenu fixe et l'inflation anticipée pourront être considérées comme des déterminants de la demande d'actifs en logements. Lorsqu'on a affaire à de telles variables inobservables on utilise souvent une procédure grossière

dépourvue de justification théorique pour incorporer leurs ersatz au modèle.

La théorie fournit rarement au prévisionniste un guide sur la manière de formuler dans le modèle une structure dynamique. Les économètres ont reconnu l'importance des structures dynamiques en modélisation et ont consenti des efforts de recherche considérables sur l'introduction de retards dans les modèles économétriques. Leurs efforts ont engendré un vaste éventail de structures de retard de remplacement, parmi lesquelles figurent les modèles à retards échelonnés géométriques et rationnels. Les justifications théoriques paraissent toutefois insuffisantes en pratique pour imposer l'emploi de structures de retard spécifiques, aussi est-il nécessaire de trouver appui sur les données.

Il apparaît que la théorie économique a encore moins à dire à propos de la structure dans les séries temporelles des termes d'erreur. On postule généralement en modélisation macroéconomique, au stade de la spécification du modèle, une structure d'erreur de type bruit blanc, quitte à tester cette supposition ultérieurement. En fait, les termes d'erreur sont ajoutés aux équations structurelles pour estimation une fois que la structure du modèle a été élaborée en termes déterministes.

La manière classique en économétrie de spécifier les erreurs dans les séries temporelles consiste à poser l'autorégression du premier ordre

$$u_t = \rho u_{t-1} + \varepsilon_t,$$

où ϵ_t est le « bruit blanc ». Aiguillonnés par les développements récents dans l'analyse des séries temporelles, les concepteurs de modèles ont, dans le cadre du modèle régressif et du SEM, envisagé des structures d'erreur autorégressives et à moyenne mobile d'ordre plus élevé.

2.2 PRÉVOIR AVEC UN MODÈLE ÉCONOMÉTRIQUE

Une fois un SEM spécifié, on peut l'estimer de manière cohérente au moyen d'une méthode à équation unique telle que celle des doubles moindres carrés (DMC) ou d'une méthode systémique telle que celle des triples moindres carrés (TMC). Si le modèle est complexe et renferme de nombreuses variables prédéterminées, de telles méthodes pourront ne pas être utilisables à cause d'un échantillonnage trop restreint. Dans cette section nous expliquons comment employer un modèle estimé pour effectuer des prévisions économétriques.

A. Prédire avec un SEM

Klein (1968) définit la prévision comme « la tentative de faire des énoncés scientifiques à propos de situations sans échantillonnage en se fondant sur des relations déterminées à partir d'observations d'échantillons ». On distingue deux types de prévisions : prévisions a priori et prévisions a posteriori. Pour effectuer des prévisions a priori, on rassemble autant d'informations pertinentes que possible et on extrapole vers un avenir inconnu avant l'événement anticipé. Pour les prévisions a posteriori, on utilise des informations pertinentes pour prédire des données déjà existantes. La prévision a posteriori est un instrument utile pour la vérification diagnostique du modèle ajusté.

On distingue deux autres types de prévisions : prévisions inconditionnelles et prévisions conditionnelles. La prévision inconditionnelle est formulée comme une prédiction non restrictive de l'événement futur. En revanche, la prévision conditionnelle est une prédiction restrictive fondée sur la supposition que certains événements exogènes ont lieu simultanément. En prévision économétrique, on effectue des prévisions conditionnelles a priori.

Supposons qu'à l'instant T, appelé origine de la prévision, des prévisions à h étapes vers l'avenir des variables considérées soient requises pour un instant futur quelconque $T + h$. Si le modèle est statique et toutes ses variables prédéterminées exogènes, et si les valeurs des variables prédéterminées à $T \pm h$ sont spécifiées, le problème prévisionnel devient un problème d'estimation des valeurs qu'auront les variables endogènes du modèle à l'instant $T + h$. On obtient ainsi la prédiction de y_{T+h} étant donné x_{T+h} en donnant aux termes d'erreur structurelle futurs la valeur de zéro et en résolvant numériquement le système non linéaire suivant :

$$f_i(\hat{y}_{T+h}, x_{T+h}; \hat{\theta}_i) = 0 \qquad i = 1, \ldots, G \qquad (2.3)$$

pour les prévisions $\hat{Y}_{T+h}$. Le vecteur θ_i en (2.3) est une estimation de θ_i, le vecteur des coefficients structurels. La plupart des propriétés statistiques des prévisions $\hat{Y}_{T+h}$ ainsi obtenues sont inconnues.

Si le système est dynamique, la prévision à étape unique vers l'avenir de Y_{T+1} à partir de l'origine T peut être obtenue en résolvant numériquement

$$f_i(\hat{y}_{T+1}, \hat{y}_T, z_{T+1}; \hat{\theta}_i) = 0, \quad i = 1, \ldots, G \tag{2.4}$$

pour $\hat{Y}_{T+1}$, où θ_i est une estimation de θ_i en (2.2), z_{T+1} représente le vecteur des valeurs futures des variables exogènes à l'instant T + 1 et où la valeur attendue de l'erreur future $u_{T+1,i}$ est donnée égale à zéro.

On peut obtenir les prévisions à h étapes vers l'avenir à partir de la formule de récurrence

$$f_i(\hat{y}_{T+h}, \hat{y}_{T+h-1}, z_{T+h}; \hat{\theta}_i) = 0, \quad i = 1, \ldots, G \tag{2.5}$$

Ainsi, les prévisions à h étapes vers le futur sont tributaires de la valeur présente Y_T des variables endogènes et des valeurs futures $z_{T+1}, \ldots, z_{T+h}$ des variables exogènes jusqu'à l'instant T+h.

B. Pratique de la prévision avec un modèle économétrique

En pratique, la prévision avec un modèle économétrique n'est pas fondée seulement sur le modèle estimé et sur les données. Evans, Haitovsky et Treys (1972) distinguent trois étapes en prévision économétrique : prédiction des valeurs futures des variables exogènes, ajustement des équations individuelles et détermination des solutions du modèle.

Les valeurs futures de la plupart des variables exogènes sont inconnues au moment de la prévision et doivent donc être elles aussi prédites. Le problème de la prévision des variables endogènes cède donc la place à celui de la prévision des variables exogènes.

Il est possible de connaître les valeurs futures de certaines variables exogènes au moment de la prédiction; c'est le cas des variables décisionnelles. Pour d'autres variables exogènes il sera peut-être nécessaire de procéder subjectivement. On pourra également tenter de les prédire en se fondant sur toute l'information disponible au moment de la prédiction, par exemple sur leur historique. Si celui-ci s'étend sur une période

suffisamment longue, on pourra avoir recours à un modèle à série temporelle.

On fait des ajustements aux équations individuelles du modèle estimé en modifiant de manière subjective les points d'intersection et/ou d'autres estimations de paramètre. Ces « facteurs correctifs » trouvent leur justification dans le fait qu'ils incorporent aux prévisions toute information et tout élément d'appréciation que le prévisionniste n'a pas explicitement introduit dans le modèle. Il pourra ainsi mettre à profit des informations fiables sur des modifications des paramètres exogènes et sur des ruptures structurelles; connaissant à l'avance des changements d'ordre législatif et institutionnel, il pourra modifier le modèle en conséquence.

Enfin, on résout le modèle (estimé et ajusté) pour en tirer des prédictions sur les variables endogènes. Si toutefois le prévisionniste a procédé à une évaluation a priori de la plage des limites probable des valeurs futures et si les prévisions obtenues comme solution du modèle se situent à l'extérieur de ces limites, le prévisionniste pourra alors les modifier, soit directement, soit en réajustant les équations structurelles ajustées et en résolvant le modèle.

La mesure dans laquelle on a recours aux ajustement subjectifs varie d'un prévisionniste à l'autre; elle dépend également de l'état de l'économie au moment où est faite la prédiction. Selon McNees (1988), le poids relatif des ajustements subjectifs en prévision économétrique aux États-Unis varie de 20 p. 100 pour Chase Econometrics à 30 p. 100 pour Wharton Economic Forecasting Associates. On peut justifier un recours si systématique aux ajustements subjectifs en faisant valoir que le prévisionniste tient compte de toute l'information disponible au moment de la prédiction et améliore ainsi la qualité de ses prévisions. La précision de la prédiction reflétera alors la qualité des informations utilisées, que celles-ci aient été incorporées directement au modèle ou par le biais des facteurs correctifs.

Jusqu'aux années 1970, le modèle économétrique à équations simultanées était peut-être l'outil prévisionnel le plus couramment employé. Le principal trait distinctif de l'approche économétrique de la prévision est que la théorie économique joue un rôle dans la spécification du modèle. Les prévisionnistes ont tendance à élaborer un modèle économétrique complexe parce qu'ils croient que les modèles simples ne sont pas réalistes et incapables de répondre à maintes questions importantes.

Les modèles économétriques présentent de nombreux avantages comme outil prévisionnel :

(1) ils constituent un instrument utile pour la transformation systématique des valeurs présentes et passées des variables endogènes et exogènes et des prédictions sur des variables exogènes en prédictions sur des variables endogènes;
(2) tant que leurs paramètres restent inchangés dans l'avenir, on peut faire des prévisions portant sur un avenir aussi lointain que nécessaire si l'on dispose de prévisions sur les variables exogènes;
(3) ils permettent de faire sur des variables liées des prévisions qui soient cohérentes entre elles puisqu'elles doivent satisfaire conjointement aux contraintes du modèle, en particulier à ses identités; et
(4) ils peuvent servir à des fins autres que la prévision; on peut par exemple les utiliser pour évaluer les conséquences de diverses politiques en simulant les chemins temporels des variables endogènes étant donné les chemins temporels spécifiques des variables exogènes.

En revanche, les modèles économétriques présentent notamment comme outil prévisionnel les inconvénients suivants :

(1) ils soumettent souvent les paramètres à des contraintes « peu crédibles », surtout lorsque le modèle est complexe;
(2) les problèmes liés aux séries temporelles tels que les perturbations structurelles auto-corrélées ne peuvent être pris en compte de manière systématique; ils sont donc souvent négligés; et
(3) on ne peut calculer les erreurs-types comme mesure de la précision des prévisions parce que celles-ci sont ajustées subjectivement au moyen de facteurs correctifs.

La prévision économétrique est entachée de nombreuses sources d'erreur prévisionnelle :

(1) la source d'erreur la plus évidente provient des termes d'erreur eux-mêmes dans les équations du modèle; il est raisonnable de supposer que les valeurs futures des termes de perturbation sont égales à zéro, leur valeur anticipée, mais leurs valeurs effectives différeront de leurs valeurs anticipées utilisées pour prédire les variables endogènes;
(2) Il est nécessaire de prédire les variables exogènes pour résoudre le modèle afin d'obtenir les prévisions des variables endogènes; les erreurs faites en

prédisant les variables exogènes en dehors du cadre du modèle économétrique se répercuteront inévitablement sur la précision des prédictions des variables endogènes;

(3) les erreurs d'échantillonnage dans l'estimation des paramètres réduiront nécessairement la précision des prévisions;

(4) le modèle lui-même est sujet à des erreurs de spécification; il n'existe pas de modèle « vrai », et

(5) l'ajustement subjectif des prévisions est une source d'erreurs même s'il est censé réduire l'ampleur des erreurs.

De nombreux chercheurs ont tenté d'évaluer l'efficacité de la prévision économétrique comparativement à celle d'autres méthodes prévisionnelles. Un procédé courant d'évaluation des modèles économétriques consiste à se représenter comme concurrents ce même modèle et un modèle ARIMA univarié ou tout autre modèle « naïf » et à comparer leur efficacité prévisionnelle directement. Les recherches du début des années 1970 indiquent que les modèles économétriques, sans ajustements subjectifs, n'ont pas fait bonne figure. Cooper (1972) a comparé les prévisions post-échantillonnage effectuées un trimestre à l'avance au moyen de modèles autorégressifs aux prévisions fournies par sept modèles macroéconomiques américains trimestriels et a observé que les premières comportaient généralement une erreur prévisionnelle inférieure à celle des secondes. En outre, Nelson (1972) a observé que les modèles ARIMA étaient plus efficaces que le modèle FRB-MIT-Penn de l'économie américaine pour la prévision a posteriori en dehors de la période d'échantillonnage.

Au fil des années, l'efficacité prévisionnelle des modèles macroéconomiques s'est accrue. Une comparaison plus récente par Longbottom et Holly (1985) de l'efficacité prévisionnelle des modèles ARIMA et du modèle de la London Business School de l'économie britannique a démontré que pour de nombreuses variables le SEM est supérieur au modèle ARIMA à la fois pour les prévisions a priori et a posteriori, bien que ce soit l'inverse pour certaines variables. De même, s'agissant de variables macroéconomiques clés, McNees (1988) a observé que les prévisions ajustées subjectivement des modèles économétriques de l'économie américaine sont plus précises que celles des modèles ARIMA univariés.

Le recours systématique aux ajustements subjectifs en prévision macroéconomique présente un grave problème lorsqu'on veut évaluer les prestations relatives de différentes méthodes de prévision. Il faut bien faire la distinction entre la capacité d'un modèle à fournir de bonnes prévisions et celle d'un prévisionniste expérimenté et bien informé muni d'un modèle bien conçu à en faire autant. Il semble, en ce qui a trait aux

prévisions à long terme, que les prévisions fournies par un prévisionniste à l'aide d'un modèle économétrique soient plus fiables que celles fournies par les simples modèles ARIMA.

2.3 LE MODÈLE DU SECTEUR DU LOGEMENT

2.3.1 THÉORIE ÉCONOMIQUE DU LOGEMENT

Dans la présente section et dans les sections suivantes nous formulons un cadre conceptuel du marché du logement au moyen duquel on peut évaluer le secteur domiciliaire dans les modèles macroéconomiques et dans les modèles nationaux du logement, en particulier à des fins prévisionnelles. Le cadre auquel nous aboutirons sera le reflet de notre compréhension des vues actuelles sur ce secteur, compréhension qui est le fruit de notre examen de la littérature récente dans le domaine de l'économie du logement.

A. La théorie de base

Toutes les recherches théoriques et empiriques en matière de logement s'articulent présentement autour d'une compréhension consensuelle des éléments de base de l'économie du logement. Les principes théoriques fondamentaux ont été formulés par Smith (1974), Arnott (1987) et Olsen (1987). Suit un bref exposé de ces principes.

i) Demande : La demande pour un logement comme « denrée » est conçue comme une demande de stock découlant de la demande de flux pour des services domiciliaires spécifiques offerts par ce stock. En ce sens, le logement est traité au moyen des méthodes standard de la théorie du capital. À cause des attributs multiples des logements contribuant à la prestation de ces services (taille, qualité, etc.), la demande de stock est généralement conçue comme la demande pour une denrée unique fournissant une synthèse de ces attributs. (Certains travaux récents ont mis l'accent sur la demande individuelle de ces attributs exprimée au moyen d'indices hédoniques.) Par ailleurs, parce que le logement est à la fois un actif et un bien de consommation, la nécessité de faire toute la lumière sur la nature de la demande oblige à tenir compte de ces deux aspects. Dans le cas des rentiers, on traite la demande comme la demande d'un bien de consommation. Il est toutefois toujours loisible d'opter pour la solution de remplacement de la propriété.

ii) Offre : Par le passé, on supposait que l'offre de logements se traduisait par de nouvelles constructions parce qu'elle constitue un flux. Mais selon la nouvelle conception stock-flux on admet que l'offre effective correspond au flux de services à partir du stock existant. Les nouvelles constructions jouent un rôle en agrandissant le stock, quoique avec retard, selon les critères habituels de profitabilité (prix anticipé relativement au coût total de la construction et du terrain). Aussi les modèles fondés sur l'ajustement du stock sont-ils monnaie courante. À cause de la fixité du parc résidentiel, l'offre de services de logement est inélastique à court terme.

Mais à long terme cette offre est beaucoup plus élastique parce que les nouvelles constructions accroissent le stock de services de logement.

iii) Nature du marché : Dans la mesure où l'on admet que le parc résidentiel est homogène, on suppose que l'écoulement par le marché est fluide et que le prix sur le marché reflète le prix d'équilibre ou de dégagement du marché.

iv) Finances : Le marché du logement doit être envisagé comme étant étroitement associé au marché financier. La raison en est que le montant des débours nécessaires pour acquérir le stock exige dans la plupart des cas le financement de la dette au moyen d'hypothèques. Aussi dans la plupart des modèles a-t-on incorporé explicitement les taux d'intérêt à la fonction de demande. Le coût et la disponibilité du crédit hypothécaire peuvent en outre se répercuter sur le marché du logement.

B. Problèmes irrésolus de la théorie de base

Divers problèmes n'ont pas été entièrement résolus par la théorie de base, notamment :

i) Demande : La manière de traiter tout le problème de l'incertitude et des attentes, eu égard en particulier aux prix et aux taux d'intérêt, reste une question controversée et importante.

ii) Offre : On reconnaît maintenant qu'il existe, en ce qui a trait à l'offre, des évolutions cruciales à court terme relatives à des ajustements quantitatifs (vacances/surpeuplement), des conversions, des démolitions, des rénovations, etc., évolutions dont il faut tenir compte à la fois pour déterminer les prix et pour comprendre les choix de statut d'occupation.

iii) Prix : L'établissement du prix du stock reste problématique parce que seule une petite fraction de celui-ci est l'objet de transactions chaque année.

iv) Le marché : À cause de la nature différenciée de la denrée considérée, s'agissant du type (logement simple, multiple), de la localisation et ainsi de suite, on a exprimé des doutes quant à savoir si on peut la traiter comme une denrée homogène dans un marché unitaire. En fait, la plupart des théoriciens l'envisagent dans le cadre d'un ensemble complexe de marchés étroitement imbriqués et interdépendants.

v) Équilibre : La nature des retards à la fois des côtés de la demande et de l'offre incite à croire que les prix d'équilibre seront rarement atteints. Cela sera vraisemblablement le cas au sein d'un marché du logement localisé (zone urbaine ou région

métropolitaine de recensement), et certainement le cas à l'échelle nationale.

vi) Politique publique : Le logement est désormais considéré comme un bien public, voire comme un droit. Aussi les décideurs à tous les niveaux y accordent-ils une attention considérable. La modicité de l'assiette fiscale des municipalités incite celles-ci à adopter des politiques de logement qui soient conçues de manière à accroître leurs revenus, souvent selon des modalités qui altèrent le marché du logement. Or ces politiques sont rarement prises en compte par les modèles de logement.

vii) Finances : La conception « traditionnelle » selon laquelle l'offre en logements est équivalente aux nouvelles constructions a des implications directes quant au traitement des marchés financiers. Ce qu'on supposait, c'est qu'il y a un lien entre certaines variables financières (taux d'intérêt, consentements, etc.) et les mises en chantier. La nature des marchés hypothécaires n'a jamais été modélisée explicitement. Voire, la complexité de ces marchés a sans doute joué un rôle dissuasif à cet égard. De plus, selon le récent paradigme de l'ajustement du stock, une telle explication causale est inadéquate. Il y a manifestement encore beaucoup de progrès théoriques à accomplir dans ce domaine.

Aucune de ces réserves n'exige de modifier la structure théorique de base du marché du logement, mais elles obligent les analystes à élaborer le modèle de base avec beaucoup plus de soin qu'il n'en faut pour la plupart des autres marchés.

C. Approches macroéconomiques versus approches de marché : quelques contrastes

Pour des raisons très pratiques, les économistes du logement ont emprunté deux voies différentes pour perfectionner le modèle de base évoqué ci-dessus.

La première de ces voies est axée essentiellement sur les prévisions à grande échelle; elle est motivée par le besoin de prédire des variables clés du secteur du logement qui sont d'intérêt primordial pour les décideurs économiques (mises en chantier, niveau d'investissement, taux d'intérêt, prix et locations, vacances, etc.). Dans la littérature sur le sujet, on s'intéresse moins au fonctionnement du marché du logement qu'aux données globales sur le logement. Au Canada, on n'a élaboré que peu de macromodèles, sinon aucun, qui soient conçus explicitement pour prévoir les données futures sur le logement au niveau macroéconomique. Ce dont on dispose, ce sont des segments de macromodèles plus compréhensifs expliquant les données futures clés sur le logement. Mais les explicitions fournies sur ces données sont en général très superficielles, car elles ne

s'appuient que sur des fondements théoriques des plus simplistes parce que l'ampleur des modèles plus vastes limite le nombre de variables et d'équations sur le secteur du logement. On met nécessairement l'accent sur des considérations très globalisantes, presque uniquement au niveau national. Enfin, les marchés du logement étant cloisonnés, les données globales utilisées dans ces modèles reflètent des situations non comparables et de déséquilibre. C'est pourquoi l'interprétation des résultats est problématique.

On rétablit l'équilibre au moyen d'une mystérieuse boîte noire d'ajustements et de transformations de stock. De plus, on traite généralement à la va-vite les ajustements complexes du marché des capitaux et des politiques publiques (aux échelles nationale et locale) qui se répercutent sur le marché du logement.

Comme nous le verrons dans la prochaine section, on a au cours des cinq dernières années réalisé d'importantes percées théoriques permettant d'affiner ce modèle de base, ce qui promet de le rendre plus utile aux analystes du logement, même à ce niveau global. L'analyse de ce qui se déroule au sein de la boîte noire est à cet égard d'intérêt particulier. On est également parvenu à une meilleure compréhension du rôle des marchés de capitaux et des politiques publiques.

La seconde voie empruntée par les analystes du logement met plutôt l'accent sur le fonctionnement du secteur du logement comme un processus de marché. La démarche est ici moins globale puisqu'elle s'articule autour de données futures spécifiques concernant par exemple le type de logement, le choix du statut d'occupation, etc. On accorde en outre beaucoup plus d'importance à la localisation du logement de même qu'à la question cruciale de l'interdépendance du marché. À cause de cette optique de détail, il s'agit généralement de modèles à équilibre à long terme. Il convient d'ajouter qu'à cause de la nature spécifique de cette optique il est difficile d'appliquer ses résultats à l'échelle nationale. Les modèles en question font un emploi beaucoup plus systématique des données. En outre, des problèmes liés à la qualité des données, à la non-comparabilité des marchés, à leur atemporalité, à leur complexité analytique, etc. rendent ces modèles peu utiles pour la prévision macroéconomique. Aussi a-t-on rarement tenté d'appliquer ces modèles de manière systématique à l'ensemble des marchés du logement.

De tels modèles, à la suite des travaux de Muth (1988), permettent néanmoins une analyse beaucoup plus détaillée du fonctionnement du marché du logement. Ils ont en particulier suscité diverses recherches sur le contenu des boîtes noire au moyen d'ajustements de stock, notamment les réparations, les démolitions, etc.

Le problème pour les prévisionnistes (et au bout du compte pour les décideurs) consiste à synthétiser ces deux voies d'analyse. Dans le cas idéal, nous souhaiterions disposer d'une approche de modélisation cohérente les incluant toutes deux, en satisfaisant à la fois aux exigences pratiques des macromodèles et aux exigences théoriques des modèles moins globaux. On a réalisé certains travaux dans cette voie, mais surtout à un niveau relativement abstrait (Carruthers (1989)). En pratique, dans la mesure où l'on cherche à la réaliser, on ne vise une telle synthèse qu'au moyen d'un bricolage empirique.

Dans la section suivante, nous examinons les développements récents (théoriques et empiriques) dans chacun de ces domaines et cherchons à entrevoir où pourrait s'articuler une synthèse plus formelle. Nous commençons par une description du cadre essentiel du macromodèle. Nous passons ensuite en revue les développements récents en macromodélisation à la lumière de ce modèle. Plus précisément, nous allons montrer comment ils accroissent la rigueur conceptuelle du modèle. Nous procéderont ensuite à une analyse des développements récents dans la sphère microéconomique qui, à notre avis, promettent d'améliorer le macromodèle.

2.3.2 CADRE DU MACROMODÈLE DU SECTEUR DU LOGEMENT

A. L'état des choses au Canada

Notre parcours de la littérature a commencé par une revue des modèles macroéconomiques existants au Canada. Nous avons eu la chance de pouvoir puiser dans des évaluations de qualité (Foot (1985) et Grady (1985)). Ce qui en ressort clairement, c'est que parce que le logement ne constitue qu'une composante modeste de ces modèles, on n'a accordé que peu d'importance à un traitement théorique rigoureux du secteur du logement. Cela constitue une sérieuse lacune des macromodèles eux-mêmes parce que le secteur du logement influe considérablement sur des variables macroéconomiques clés à cause de la contribution du logement à l'instabilité macroéconomique et de la jonction cruciale entre les dimensions monétaire et réelle de l'économie qui s'opère au sein de ce secteur.

Tous les modèles du secteur du logement sont fondés essentiellement sur le même paradigme du logement que celui décrit dans la section précédente, à savoir une demande de stock se répercutant sur l'offre fixe existante, d'où fixation du prix, avec ajustement consécutif de l'offre au prix (principalement sous forme de nouvelles constructions).

En se fondant sur ce paradigme, Foot conclut qu'on n'a pas assis le secteur du logement dans les macromodèles canadiens

actuels sur des fondations microéconomiques solides. En particulier, ces modèles négligent l'analyse du parc résidentiel et n'analysent pas de manière adéquate les taux d'intérêt, le cloisonnement du marché, le marché de la revente, l'indice des prix des logements et les retards importants qui caractérisent ce secteur.

Il est utile à ce point de procéder à une évaluation plus détaillée du secteur du logement dans les macromodèles canadiens. Grady en fournit d'excellents résumés (pp. 253-293). Dans la quasi-totalité de ces modèles, on s'attache presque exclusivement à estimer le niveau d'investissement en construction résidentielle. Grady divise les modèles en deux grandes catégories :

(A) Les mises en chantier sont liées essentiellement aux consentements hypothécaires. Il convient de noter, toutefois, que dans cette approche la relation de causalité dans le marché du logement est erronée : normalement, les mises en chantier précèdent les consentements hypothécaires.

(1) Candide 2.0 (Conseil économique du Canada) : la demande (qui est fonction des consentements hypothécaires, des variables démographiques et du coût de la propriété relativement à celui de la location) et l'offre (modifications quant aux vacances, politiques gouvernementales) déterminent les mises en chantier simples et multiples.
(2) RDX2 (Banque du Canada) : la demande est fonction de la disponibilité du financement hypothécaire.
(3) FOCUS (Université de Toronto) : la demande est fonction de la disponibilité du financement hypothécaire.

(B) Les mises en chantier sont modélisées directement :

(1) En fonction uniquement de considérations liées à la demande :
   i) TIM (Informetrica) : modèle à ajustement du stock commandé par la demande;
   ii) DRI (Data Resources of Canada) : modèle à ajustement du stock commandé par la demande.

(2) En fonction à la fois de la demande et de l'offre :
   i) RDXF (Banque du Canada) : le prix (MLS) est déterminé par l'écart entre la demande et l'offre (mises en chantier);
   ii) QFS (Ministère des Finances) : semblable au RDXF;
   iii) CHASE (Chase Econometrics) : la demande (le pouvoir d'achat effectif réel et les taux hypothécaires) et l'offre (prix de vente relativement aux coûts de construction)

déterminent le nombre de mises en chantier par habitant;

iv) MTFM (Conference Board of Canada) : demande, offre (profitabilité des mises en chantier) et assistance gouvernementale.

La conclusion critique de Grady de son examen exhaustif mérite d'être notée :

> Le logement est un secteur problématique. ... Il se peut fort bien que la seule solution valable soit de faire table rase et d'élaborer un modèle structurel du marché du logement qui décrive fidèlement le comportement de tous les agents et qui spécifie clairement la demande et l'offre dans tous les sous-marchés (p. 237).

Récemment, le ministère des Finances a élaboré un nouveau macromodèle, le modèle économique et fiscal canadien (MEFC), que Stockes (1987) décrit en détail. Quoique fondé sur le même paradigme de base stock-flux, le modèle incorpore dans une large mesure les améliorations clés décrites dans la littérature. Notamment : la dimension offre est développée de manière à inclure une fonction de production spécifique pour les nouveaux logements; les liens cruciaux entre les dimensions financières et réelles des marchés sont explicités; les attentes sont incorporées explicitement et les issues à la fois à court terme (déséquilibre) et à long terme (équilibre) sont traitées.

La structure essentielle du modèle est la suivante :

(1) les mises en chantier sont tributaires de l'offre de nouveaux logements et dépendent de la profitabilité, elle-même fonction du prix de vente et des coûts de construction anticipés;
(2) le stock tient compte des démolitions et des achèvements (fonction des mises en chantier plus le temps); et
(3) les coûts de construction varient alors en fonction des nouveaux logements (mises en chantier présentes et passées) et de coûts additionnels.

Cela constitue manifestement un modèle beaucoup plus satisfaisant du comportement du marché du logement.

Notre propre analyse indique que le traitement du côté offre, surtout pour les nouveaux logements, reste particulièrement inadéquat dans la plupart des modèles (sauf le MEFC), parce qu'on néglige des éléments clés tels que les fonctions de production, des éléments de coût essentiels et les changements technologiques. En outre, le processus d'ajustement

du stock est traité de façon simpliste dans presque tous les modèles.

Nous n'avons trouvé dans aucun modèle un sous-marché satisfaisant pour l'aspect financier. Plus précisément, en dépit de son importance, le comportement du marché hypothécaire n'a pas été introduit comme variable endogène dans les modèles de logement. Les raisons en sont claires. Il se peut fort bien qu'il soit tout simplement impossible de spécifier au moyen de quelques équations fondamentales ce qui est en réalité un marché très complexe dominé par des institutions oligopolistiques gérant à la fois les prix et les stocks, et cela a peut-être découragé les concepteurs de modèles de consacrer beaucoup d'efforts à ce sous-secteur. (Hatch (1975) a fourni une description quelque peu périmée mais utile du marché hypothécaire canadien.)

La seule autre tentative de modéliser le secteur du logement selon une optique globale est celle de Clayton (1987). Il s'agit en réalité essentiellement d'un exercice prévisionnel, et le « modèle » est plutôt mécanique. Il se compose des éléments suivants :

(1) la dérivation des projections sur les ménages est tributaire des projections sur la population; les ménages sont décrits selon l'âge, le type, le statut d'occupation et enfin le type de logement; on fournit également des projections régionales; la méthode d'analyse est la méthode des cohortes, et les variables sur le secteur du logement du modèle ne sont pas endogènes; les tendances sont plutôt décrites de manière subjective, ce que reconnaît l'auteur; cela vaut également pour le choix du statut d'occupation, qui n'est donc pas une fonction d'issues de marché explicitement modélisées, de même que pour le type de logement;
(2) la demande de remplacement est évaluée selon le statut d'occupation et le type de logement; les nombres obtenus sont très approximatifs faute de données fiables;
(3) estimation du taux d'inoccupation; il ne s'agit encore que d'estimations grossières; et
(4) la projection des besoins en nouveaux logements selon le statut d'occupation et le type de logement est fondée sur la nouvelle demande (1), la demande de remplacement (2) et l'évolution du taux d'inoccupation (3).

Manifestement, cette tentative a peu à voir avec la prévision macroéconomique. Elle ne fait usage de presque aucune des variables économiques traditionnelles, telles que prix, revenus et taux hypothécaires. Chose significative, il n'y a aucune référence dans le texte à quelque macromodèle ni même à la

littérature théorique. Étant donné qu'on y adopte une optique à long terme, on pourrait avancer que les variables macroéconomiques traditionnelles seraient ici peu utiles parce qu'elles se limitent à des horizons temporels plus rapprochés. Mais il est difficile d'admettre que les variables économiques ne jouent aucun rôle à long terme et, plus précisément, qu'on peut les incorporer implicitement par voie subjective au moyen de facteurs de pondération et ainsi de suite.

En revanche, on pourrait considérer une telle tentative comme un commentaire sur l'utilité des modèles actuels à des fins prévisionnelles pratiques. Quoi qu'il en soit, cette étude est sans grande utilité pour expliquer le fonctionnement du secteur du logement, que ce soit en termes macroéconomiques ou en termes de marché.

B. Autres développements des modèles de prévision macroéconomiques

Mis à part le MEFC, on ne voit guère de percée importante en macromodélisation du secteur du logement au Canada. Mais dans d'autres pays on a mis au point des modèles prévisionnels explicites pour le logement de niveau largement supérieur à ce qui semble avoir été accompli au Canada.

Une étude australienne innovatrice de Williams (1984) met à profit certaines des percées décrites dans la littérature. Un modèle à moyen terme à 13 équations tient compte explicitement des liens entre choix de portefeuille, prix du logement et constructions nouvelles. Un modèle généralisé d'ajustement de l'actif fournit une bonne explication du prix d'actif du parc existant (et des terrains attenants). Les constructions nouvelles y sont déterminées par les prix d'actif et le calcul de l'élasticité du nombre total de mises en chantier eu égard au prix d'actif donne une valeur de 1,4, ce qui a d'importantes implications quant aux politiques à adopter sur le contrôle du prix des logements. Mais diverses lacunes graves dans les données empêchent de parvenir à des estimations tout à fait satisfaisantes. Enfin, fait significatif, l'auteur est d'avis que :

> On peut s'attendre à ce que le modèle élaboré dans cet article fournisse de meilleurs résultats si l'on procède aux estimations selon les régions. Les limites de validité du concept de valeur moyenne des logements en Australie sont évidentes (p. 153).

Divers autres modèles orientés essentiellement vers le marché (par opposition à ceux orientés vers la prévision), mais comportant des implications pour la macromodélisation, sont passés en revue dans une section ultérieure. Mais il convient

d'accorder une attention toute particulière à l'étude de prévision macroéconomique de Goodman et Gabriel (1987), car elle est peut-être bien à la pointe du progrès pour de tels modèles.

Dans cette étude, les auteurs s'attachent à déterminer les causes précises de la piètre qualité des prévisions sur l'activité de construction résidentielle aux États-Unis dans les années 1980. Ils soulignent l'importance des changements structurels qui s'y sont produits et le fait qu'ils ont été omis dans la plupart des modèles. Ils poursuivent leur étude en analysant les sources d'influence clés sur la construction résidentielle en tenant compte de ces changements structurels.

Voici leurs conclusions résumées :

(1) les nouveaux instruments de financement, tels que les ARM, n'ont eu qu'un impact modeste;
(2) l'abordabilité et les attentes ont eu un impact favorable;
(3) des changements dans la structure des impôts fédéraux ont eu un impact favorable;
(4) le financement exonéré d'impôts du crédit hypothécaire par voie d'obligations a eu un impact favorable mais il a dans une large mesure pris la place des hypothèques traditionnelles;
(5) les changements démographiques, surtout la création réduite de ménages au début des années 1980, ont eu un impact défavorable, mais cette tendance s'est inversée lorsque la création de ménages s'est de nouveau accrue vers la fin de la décennie; et
(6) des changements dans les tendances du développement économique régional qui ont suscité un accroissement des coûts dans le Nord-est se sont traduits par un déclin de l'activité de construction.

Les auteurs ont également constaté que les variables explicatives traditionnelles ont joué un rôle important, en particulier la croissance du revenu et les taux d'intérêt. Si les prévisions ont été peu précises, c'est qu'on a mal prédit l'évolution de ces variables.

C. Un paradigme modifié

Dans la présente section, nous allons présenter une version modifiée du modèle traditionnel (tel que résumé par Foot (1985)) de manière à incorporer plusieurs des idées nouvelles exposées dans la littérature, dont certaines ont déjà été introduites dans le modèle MEFC et les modèles étrangers.

On trouvera à la Figure 2.1 le cadre conceptuel modifié[1]. Voici les éléments essentiels du paradigme modifié :

(1) la demande de services de logement est une fonction des prix, du revenu (fixe), des facteurs démographiques, des variables financières (taux et conditions hypothécaires), des variables d'ordre politique et des facteurs macroéconomiques clés (inflation, emploi));

(2) la demande de services de logement se traduit par la demande du parc (stock) de logements requis pour fournir ces services;

(3) l'offre de stock à court terme correspond au stock net existant; celui-ci comprend le stock brut existant, plus les aménagements (remises en état) moins les suppressions (démolitions, dépréciation); l'investissement nouveau n'est pas équivalent a la fourniture immédiate de services de logement, mais constitue une réaction retardée sous forme d'augmentation du stock à des signaux provenant du marché du logement à court terme; le cheminement par lequel l'investissement nouveau se traduit en stock de logement réel est le suivant : consentement, mises en chantier, achèvements;

(4) l'investissement nouveau en logements agit directement sur les variables macroéconomiques (demande globale) et sur les variables monétaires à cause de son ampleur.

(5) l'interaction entre demande et offre de stock de logement constitue le processus d'ajustement du stock à court terme; le grippage des mécanismes d'ajustement engendre l'inoccupation, les ménages réagissant par la mobilité; il en résulte une filtration du stock par les demandeurs, dont certains changent de statut d'occupation, le propriétaire devenant locataire et inversement; prix et loyers varient en réponse à ces ajustements, mais l'équilibre est rarement atteint; et

(6) le marché hypothécaire est traité de manière endogène pour qu'il soit possible de (tenter de) prédire avec un peu plus de rigueur le rôle des variables financières clés.

Ce paradigme modifié diffère du paradigme traditionnel surtout par ce sur quoi il insiste.

---

1 Les figures et les tableaux sont présentés à la fin de chaque chapitre.

(1) Il propose de mettre beaucoup plus l'accent sur le processus d'ajustement du stock et sur les diverses transactions qui sous-tendent ce processus. Cela est essentiel si nous voulons lier les développements dans les marchés du logement existant à l'activité de construction nouvelle, jusqu'ici l'objet d'une plus grande attention. À cause de leur interdépendance dans le modèle stock-flux, ce lien est vital. Une telle optique oblige à insister plus sur l'estimation des variables clés dans le processus d'ajustement du stock, notamment le taux d'inoccupation, l'activité de rénovation et les suppressions. Elle fait également ressortir la nature des transactions passées et présentes et l'évolution des prix relatives aux habitations existantes.

(2) Notre cadre conceptuel met en relief la position unique des marchés de capitaux. En effet, nous souhaitons encourager l'incorporation aux nouveaux modèles des nouvelles idées sur le logement comme actif et sur l'impact sur la demande de logement d'influences transmises par le biais des marchés de capitaux, par exemple celles exercées par l'inflation, les changements d'ordre fiscal et ainsi de suite. La structure très fluctuante des marchés de capitaux a fait varier sensiblement l'impact de ces variables; aussi les modèles du secteur du logement, qui sont si dépendants des variables financières, doivent-ils pouvoir refléter des variations.

(3) Enfin, la littérature très abondante sur le choix du statut d'occupation et sur ses implications politiques témoigne des tentatives d'introduire ce choix directement dans le cadre du modèle de logement.

C'est à la lumière de ce cadre conceptuel amélioré que nous avons entrepris notre revue de la littérature dans la section suivante. Nous allons le modifier encore après cette revue détaillée de la littérature microéconomique.

Sur le plan conceptuel, le principal problème que présente le secteur du logement dans tous les modèles macroéconomiques disponibles au Canada est que le lien avec les fondements microéconomiques de l'économie du logement y est dans le meilleur des cas traité de manière inadéquate (Grady (1985), pp. 280-82). De plus, ces modèles sont dépourvus d'analyses de niveau inférieur au niveau national. Ce problème revêt un caractère aigu dans le secteur du logement parce que les divers sous-marchés se trouveront nécessairement à court terme à des stades différents de déséquilibre car il n'existe pas de mécanisme commercial ou de flux factoriel rapide pour les ajustements à court terme entre marchés. Il est moins problématique de traiter

ces marchés globalement et à long terme à cause de la mobilité des personnes et du capital fongible. Mais ces modèles sont par nature applicables à court terme, c'est pourquoi le traitement du secteur du logement laisse nécessairement à désirer.

2.3.3 CADRE DU MICROMODÈLE DU MARCHÉ

Smith, Rosen et Fallis (1988) ont bien décrit le niveau général de compréhension des marchés du logement et résumé l'état actuel des choses en matière de théorie et de pratique. Mais ils ont surtout concentré leur attention sur la recherche orientée vers le marché, n'évoquant qu'au passage les macromodèles.

Leur objectif était de voir comment les récentes recherches ont abordé les problèmes particuliers qui sont propres au secteur atypique qu'est le secteur du logement. Ils ont appuyé sur le fait qu'ils n'ont pas pu subsumer toutes ces recherches diverses sous un modèle global unique. Ils ont certes observé que les modèles se recoupent de plus en plus, ce qui permet de supposer que tôt ou tard il y aura convergence dans la compréhension de la nature des marchés du logement. Mais il n'en reste pas moins, et cela est important, qu'il est très peu probable à ce stade qu'on puisse élaborer à des fins de prévision macroéconomique un modèle général du secteur du logement qui parvienne à décrire adéquatement les mécanismes fondamentaux du marché du logement. Il est plutôt à prévoir que les macromodèles resteront longtemps superficiels et insatisfaisants sur le plan théorique.

On peut résumer brièvement comme suit leurs principales observations :

(1) il y a eu déplacement à partir d'une demande et d'une offre pour des services de logement vers une demande et une offre de stock;
(2) l'accent mis sur le stock permet de mieux faire ressortir le rôle des ajustements de stock (entretien, filtration) et de rendre compte explicitement des situations de déséquilibre (inoccupation), de l'influence du stock sur les prix et, par voie de conséquence, sur les constructions nouvelles;
(3) on a en outre insisté de plus en plus sur l'hétérogénéité du stock, en distinguant ses caractéristiques qu'on a analysées au moyen d'indices hédoniques;
(4) on a tenu compte explicitement du fait que le choix d'un logement s'étend sur une période prolongée;
(5) on a pris en considération la demande d'un logement à la fois en tant qu'investissement et que bien de consommation, d'où la nécessité d'une analyse de portefeuille; et

(6) enfin, on a poursuivi les recherches sur les modèles d'équilibre généraux, mais ceux-ci restent trop théoriques, abstraits et difficiles à résoudre; en d'autres termes, ils ne se prêtent pas encore à la macromodélisation.

Eu égard aux considérations macroéconomiques, ils ont signalé divers domaines où des progrès ont été accomplis. Un de ceux-ci concerne l'impact des variables financières (sur l'instabilité du secteur de la construction et donc sur les coûts), l'accent étant mis sur les facteurs institutionnels et les lacunes de l'information, outre les taux d'intérêt. On a également beaucoup étudié l'impact de l'inflation, directement sur la demande de logement et indirectement sur le marché hypothécaire (le problème du déséquilibre).

Notre propre examen approfondi des travaux accomplis depuis 1985 confirme généralement cette évaluation. Nous avons toutefois choisi de résumer la littérature à la lumière de notre cadre conceptuel modifié, pour satisfaire tout particulièrement aux besoins des concepteurs de modèles et des prévisionnistes.

Côté demande, le débat s'est poursuivi sur la valeur juste du coefficient d'élasticité. Harmon (1988) observe que même avec des mesures différentes du revenu fixe, les résultats obtenus diffèrent peu. La signification de ces coefficients est toutefois controversée à cause de difficultés relatives aux données groupées (Rosenthal (1986) ou à l'insertion combinée dans l'équation de demande de décisions multiples telles que la formation de ménages, le choix du statut d'occupation et la quantité demandée. Denton, Robb et Spencer (1986) ont développé un modèle de dépense qui tient compte du statut d'occupation. Dynarski et Sheffrin (1985) affirment avec insistance qu'il faut également tenir compte du revenu transitoire, car on a constaté qu'il joue un rôle clé dans le choix du statut d'occupation et dans la mobilité des propriétaires. En outre, la polémique demande de stock versus demande de flux se perpétue. Dans une récente étude empirique, Glennon (1989) tente d'intégrer cette problématique au schéma logement comme bien de consommation plus investissement, en associant la demande de stock au dernier et la demande de flux au premier. Muth (1986) plaide éloquemment par ailleurs en faveur des attentes adaptatives au détriment des attentes rationnelles.

Une grande partie des travaux récents sur la demande de logement a été consacrée aux variables démographiques. Une idée importante à cet égard est l'argument de Mutchler et Krivo (1989) selon lequel les variables démographiques sont en fait dépendantes du stock de logement, surtout par le biais de la formation de ménages et inversement, de sorte que les considérations démographiques devraient être traitées de manière endogène dans les modèles de logement. Il s'agit là,

certainement, d'une considération importante s'agissant des modèles à long terme. Il semble également ressortir de nombreuses études que les demandes de logement des personnes âgées et des jeunes méritent d'être analysées beaucoup plus profondément qu'on ne l'a fait jusqu'ici, étant donné leur part croissante du marché du logement.

Les conséquences de l'inflation sur la demande de logement, et surtout sur les estimations de l'élasticité du revenu, restent une question pendante, en particulier depuis le développement des modèles à choix de portefeuille. Pelser et Smith (1985) ont élaboré un modèle décrivant également l'impact de l'inflation sur le choix du statut d'occupation en se fondant sur un modèle du coût pour l'utilisateur. On a réalisé étonnamment peu de travaux sur l'effet sur la demande des impôts autres que les prélèvements à l'échelle nationale. Compte tenu de la prolifération des prélèvements à l'échelle locale par les commissions scolaires et les municipalités, ce domaine devra être étudié soigneusement, car les montants en jeu sont considérables tout comme les effets sur le logement.

Le côté <u>offre</u> du marché du logement continue à recevoir beaucoup moins d'attention, conséquence peut-être du traitement insuffisant du côté offre dans de nombreux macromodèles. La fonction production ou coût pour les nouvelles constructions est tout à fait négligée. En dépit de l'importance croissante du coût des terrains dans le coût des habitations, l'influence des terrains reste invisible dans les macromodèles. De même, d'autres facteurs, tels que le coût de la main-d'oeuvre et des matériaux, l'infrastructure, la productivité, etc., ne sont qu'effleurés. On a accordé beaucoup plus d'attention aux nouvelles constructions résidentielles et aux débours pour investissement. Dans une étude récente, Topel et Rosen (1988) amorcent l'étude de ces questions au moyen d'un modèle fondé sur la fixation dynamique du prix selon le coût marginal. Nous avons relevé peu de travaux nouveaux sur l'importante question macroéconomique de l'instabilité économique et son influence sur l'investissement domiciliaire, en dépit du caractère volatil de l'économie depuis le début des années 1970.

Conséquence de la prise de conscience croissante que le flux de services de logement dépend de la taille et de la composition du stock net disponible de logements, de nombreux travaux de recherche ont été entrepris pour tenter de définir et de mesurer le volume du <u>stock net disponible au sein du stock brut existant</u>. Hendershott et Smith (1988) ont élaboré un cadre général en soulignant le rôle de l'ajustement du stock comme outil d'analyse d'ajouts au parc résidentiel jusque-là non comptabilisés. On a beaucoup analysé en détail les ajouts (autres que la construction) au stock brut et les retraits de ce même stock. On a étudié le rôle des rénovations et des remises en état, surtout à partir d'une perspective théorique, mais également d'un point

de vue empirique (Boehm et Ihlanfeldt (1986). S'agissant des retraits, on a concentré le regard sur l'amélioration de l'estimation de la dépréciation (Malpezzi, Ozanne et Thibodeau (1987)) et sur la suppression des logements.

Étant donné l'utilisation croissante du modèle stock-flux, il est devenu impératif de mieux comprendre les mécanismes d'affectation du stock. Aussi a-t-on abordé l'étude du processus d'ajustement du stock selon diverses perspectives. Mais on y considère le processus dans sa totalité en étudiant les modèles d'ajustement dans le marché du logement. Dans un récent modèle à long terme du marché du logement américain, Gahvari (1986) a calculé une élasticité intertemporelle explicite de substitution entre le logement et les autres biens de consommation. Turner et Struyk (1985) emploient l'Urban Institute Market Simulation Model, dont le cadre prévisionnel s'étend sur dix ans, pour déterminer par projection les effets à long terme des politiques aux États-Unis. Un modèle explicite du marché de la revente a été élaboré par Rosen et Smith (1986), dans lequel on analyse en particulier l'effet sur les prix de la demande et de l'offre de stock.

On s'est moins attaché à décrire le comportement des acteurs dans le marché du logement comme composante du processus d'ajustement. Le comportement réel des vendeurs de stock a toutefois été analysé explicitement en Suède (Aguilar et Sandelin (1984)). De manière plus générale, on a fait usage dans les études sur les transactions de modèles de recherches et d'enchères. On a examiné le marché du logement en tant que marché contractuel, fournissant ainsi de nouveaux aperçus sur le processus d'ajustement. Le rôle des intermédiaires dans ces transactions a également été passé au crible (cf. Jud et Frew (1986)).

À cause des retards dans les ajustements du marché du logement, il a été nécessaire de considérer explicitement dans plusieurs modèles d'ajustement du stock les ajustements de quantité de même que ceux de prix. Une variable clé à cet égard est le taux d'inoccupation. Emmi (1984) a proposé un moyen innovateur de traiter cette question, à savoir la modélisation du transfert des logements vacants. Bien entendu, les ménages procèdent également à des ajustements en changeant de résidence ou, plus généralement, par leur mobilité. En Amérique du Nord on a réalisé peu de travaux traitant explicitement de ce sujet, même si Quigley (1987) a analysé l'impact des variables financières sur la mobilité. Dans certaines recherches menées en Hollande on a souligné l'importance de la mobilité, laquelle sera sans aucun doute étudiée ici de manière plus systématique à l'avenir.

La filtration constitue un procédé très utile pour associer les logements vacants et les ménages mobiles dans le cadre du processus d'ajustement du marché du logement. Baer et Williamson

(1988) ont développé un modèle présentant un cadre de recherche unificateur. En outre, ils distinguent des domaines méritant d'être l'objet de recherches futures sur ces questions. Wicher et Thibodeau (1988) soulignent l'importance de la filtration pour l'adoption de politiques.

Un facteur important ayant de profondes répercussions sur l'ajustement du stock est le choix du statut d'occupation. À des fins décisionnelles, la distinction entre la situation des propriétaires et celle des locataires est en effet d'importance capitale. Aussi la complémentarité de ces deux points de vue a-t-elle suscité l'apparition d'une volumineuse littérature sur le choix du statut d'occupation. Krumm (1987) démontre de manière plutôt convaincante que de tels choix doivent être envisagés dans la durée, et que les conclusions relatives à une période donnée peuvent être trompeuses. Les variables pertinentes sur le revenu et la fortune influant sur le choix du statut d'occupation ont été largement débattues. Enfin, de nombreuses études ont signalé que le phénomène des condominiums constitue une étape intermédiaire utile entre les solutions traditionnelles de la location et de la propriété.

On continue à explorer le rôle des finances. Dans la plupart des recherches empiriques on a examiné l'influence des variables financières clés plutôt que le fonctionnement du marché hypothécaire lui-même. Ainsi, Muth (1986) a procédé à des estimations de l'élasticité de l'offre de prêts hypothécaires et de ses répercussions sur le logement. On a constaté que le rationnement du crédit est important à court terme mais pas à long terme. Nous avons noté qu'on a peu progressé depuis les travaux descriptifs des années 1970 (Andersen et Ostas 1977)) ou les travaux théorico-empiriques de cette période (par exemple, Smith et Sparks (1970)). Ce dernier travail est important parce qu'on y articule un modèle du marché hypothécaire dans lequel on associe les taux hypothécaires directement aux taux LNH et aux taux obligataires et indirectement aux ajustements de portefeuille des institutions financières. Une étude plus récente par Jaffee et Rosen (1979) sondant l'efficacité des politiques de stabilisation est également fondée sur un modèle d'intérêt particulier. Plutôt que d'utiliser un modèle à équation unique et à variable « prix » unique, ils font usage à la fois du taux contractuel et d'autres termes hypothécaires. Ils comparent de plus les processus d'ajustement partiels et instantanés et observent que les premiers sont plus efficaces.

Il existe donc des éléments à partir desquels un modèle complet du marché hypothécaire pourrait être élaboré et inséré dans un modèle de logement plus vaste. Mais à ce jour rien ne nous indique que cela ait été accompli. Le rôle fluctuant des variables financières a été analysé dans les articles importants de Friedman (1989) et Kahn (1989) qui soulignent l'importance des changements d'origine institutionnelle et autre des variables monétaires, compliquant d'autant la tâche du prévisionniste.

Goodwin (1986) a étudié l'étroite interdépendance entre les marchés financier et domiciliaire, les deux tendant à être en déséquilibre.

Des issues de marché du logement sont analysées dans de nombreux articles importants. Hendershott (1988) attribue l'accroissement du prix des maisons dans les années 1980 principalement à la baisse de la productivité et l'accroissement de la propriété à l'augmentation du revenu. On a constaté que l'impact du taux d'inoccupation a été moindre que ne l'avait cru auparavant Gilderbloom (1986).

À partir d'une perspective plus globale, où la structure entière des prix des logements dans le pays est essentielle, il importe de comprendre dans les faits comment les prix dans un marché donné influent sur ceux d'un autre. Rosenthal (1986) analyse l'interdépendance spatiale des marchés du logement et conclut qu'à long terme, à cause de la complexité des interactions, cette interdépendance est peu marquée. Si cela devait se confirmer, alors il serait tout à fait impératif de traiter les marchés individuellement pour les prévisions à court terme. Manning (1986) a élaboré un modèle qui explique environ 70 p. 100 des variations entre villes du prix des habitations.

2.4 LES DONNÉES

La capacité à élaborer un modèle économétrique du secteur du logement et du marché hypothécaire dépend non seulement de la théorie économique à laquelle on a recours pour le spécifier mais de la disponibilité et de la qualité des données à utiliser. C'est pourquoi notre examen comprend un examen des principales sources de données disponibles pour la modélisation économétrique de ce secteur au niveau national.

Nous avons divisé les variables du secteur du logement et du marché hypothécaire en cinq catégories selon le domaine considéré : demande de stock, offre de nouveaux logements, offre de stock, dégagement du marché et finances. Nous avons accordé une attention toute particulière à la fréquence de la disponibilité des données. Nous nous sommes également intéressés à la qualité des données. Pour déterminer cette dernière, nous avons consulté à la fois les fournisseurs de données et les utilisateurs, surtout ceux qui s'emploient à élaborer et à exploiter les principaux modèles macroéconomiques canadiens.

A. Demande de stock

Les données sur le prix des logements existants sont disponibles mensuellement, trimestriellement et annuellement auprès du Service interagences (SIA ou MLS) de l'Association canadienne de l'immeuble. Quoique ces données soient abondantes,

la proportion des logements inscrits au SIA n'en varie pas moins de manière cyclique. Nous n'avons pas été en mesure de déterminer jusqu'à quel point ces fluctuations faussent les estimations des prix du marché.

S'agissant des loyers il existe des sources principales de données; l'indice des prix à la consommation (mensuel et annuel) et les rapports d'enquête de la SCHL (semi-annuels). Le premier assure au cours du temps un meilleur contrôle de la qualité mais ne dispose pas d'échantillons de grande taille comme les seconds. Certains sont d'avis que les rapports d'enquête rendent compte plus fidèlement des loyers du marché (par opposition aux simples mouvements de prix). Il convient en outre de noter qu'on admet généralement que la composante loyer de l'IPC est biaisée vers le bas.

Les données sur les avoirs des ménages sont de piètre qualité. Elles ne sont pas obtenues de manière régulière et leur couverture est incomplète. Les sources sont Statistique Canada et l'EDF.

Est tout à fait surprenante la mauvaise qualité des données sur le nombre de ménages, égale par définition au stock de logements occupés. On peut se procurer des données annuelles HFE/ERMEM et de la Division des sciences, de la technologie et du capital action de Statistique Canada, bien qu'elles soient considérées de qualité très douteuse par les utilisateurs. Les premières proviennent d'estimations de la population totale et de la taille moyenne des ménages; pour obtenir les secondes, on estime d'abord le stock de logement net et, en tenant compte ensuite du taux d'inoccupation, on calcule ensuite le stock occupé. Ces deux sources de données fournissent donc des estimations indépendantes de la même variable.

Les données sur l'impôt foncier des ménages fournies par les Comptes provinciaux de Statistique Canada paraissent satisfaisantes.

B. Offre de nouveaux logements

L'offre de nouveaux logements dépend très étroitement de leur profitabilité pour les propriétaires et les constructeurs.

Statistique Canada fournit des indices du prix des nouvelles habitations (mensuels, trimestriels et annuels). La SCHL fournit également des données mensuelles et trimestrielles sur le prix de vente moyen à l'unité de maisons individuelles récemment achevées. Ces deux sources offrent un point de départ pour la collecte de données sur les prix.

Eu égard au coût des terrains, de la main-d'oeuvre et des matériaux, le Recensement de la construction de Statistique

Canada fournit des données annuelles, mais leur qualité n'a pu être vérifiée. On peut en outre se procurer de Statistique Canada des indices mensuels et annuels des coûts de la main-d'oeuvre et des matériaux. Les indices du coût de la main-d'oeuvre sont fondés sur les contrats salariaux des syndicats, lesquels sous-estiment manifestement le coût de la main-d'oeuvre. Quant aux indices sur le coût des matériaux, ils ne reflètent pas les coûts réels des constructeurs parce que ceux-ci ne payent pas tous les mêmes prix. Les indices de coût fournis par Statistique Canada semblent être de qualité douteuse.

Les données annuelles sur le barème de l'impôt pour l'industrie de la construction paraissent adéquates, mais l'obtention de données sur le seul secteur résidentiel est plus coûteuse. Les données annuelles sur les indemnités pour consommation du capital ne concernent que l'ensemble de l'industrie de la construction, aussi sont-elles trop globales pour être utiles pour l'analyse de la seule sous-composante résidentielle. La qualité des estimations annuelles sur les profits est incertaine, bien qu'on nous ait conseillé de les manier avec beaucoup de prudence. Seul le Recensement de la construction apporte des informations sur la sous-composante résidentielle.

Si les données sur les coûts d'intrant sont de mauvaise qualité, les mesures de la construction d'habitations sont en revanche de très haute qualité. On peut se procurer à la SCHL des données mensuelles, trimestrielles et annuelles sur les mises en chantier et les achèvements d'habitation. Cet organisme fournit également des données mensuelles et annuelles sur les dépenses pour les nouvelles constructions. Il publie des renseignements sur les coûts de construction estimés des constructions nouvelles et existantes, mais uniquement pour celles qui sont financées dans le cadre de la LNH. Étant donné que la valeur des terrains n'est pas incluse, ces données sous-estiment lès véritables frais de construction. La SCHL publie en outre des chiffres annuels sur le parc résidentiel construit dans le cadre des programmes publics.

C. Offre de stock

Les utilisateurs des données annuelles sur le stock brut de Statistique Canada considèrent que ces données présentent de très sérieux problèmes, à telle enseigne qu'on a fait valoir qu'à cause de cette lacune il est présentement impossible d'élaborer un modèle économétrique du secteur du logement. Il convient également de noter que les données du recensement de la construction de Statistique Canada sont publiées avec un retard de deux ans.

Les données annuelles sur les retraits sont tributaires des rapports sur les démolitions de Statistique Canada, lesquels paraissent extrêmement suspects à cause de problèmes relatifs à la collecte des données. Surviennent également des problèmes de cohérence suscités par le taux de couverture croissant au cours des années. On considère généralement que les données de l'enquête annuelle sur les rénovations sont de piètre qualité. On peut se procurer auprès de Statistique Canada des données trimestrielles et annuelles sur la dépréciation au coût de remplacement, données qui semblent adéquates. Les données annuelles sur le stock net souffrent des mêmes lacunes que celles sur le stock brut.

D. Variables sur le dégagement du marché

Les données annuelles et semi-annuelles de la SCHL sur le taux d'inoccupation sont biaisées vers le bas à cause d'une définition très restrictive. On peut se procurer du SIA des données mensuelles, trimestrielles et annuelles sur les ventes. Elles sont sujettes à un biais systématique, comme nous l'avons signalé plus haut à propos des prix des logements existants. On estime que l'autre source de données, celles de Teela, sont compilées négligemment, aussi ont-elles mauvaise réputation auprès des utilisateurs et des fournisseurs de données.

Les données annuelles sur le statut d'occupation (HFE) et sur le type de logement (SCHL, Recensement de la construction) sont en général tout à fait acceptables.

E. Finances

Des données sur les taux hypothécaires (SCHL) sont disponibles mensuellement et trimestriellement, sur les consentements hypothécaires (SCHL) mensuellement, trimestriellement et annuellement, sur la disponibilité du crédit (Banque du Canada) mensuellement et annuellement et sur les frais de possession (ERMEM) annuellement. Elles sont toutes considérées de très bonne qualité.

F. Conclusion

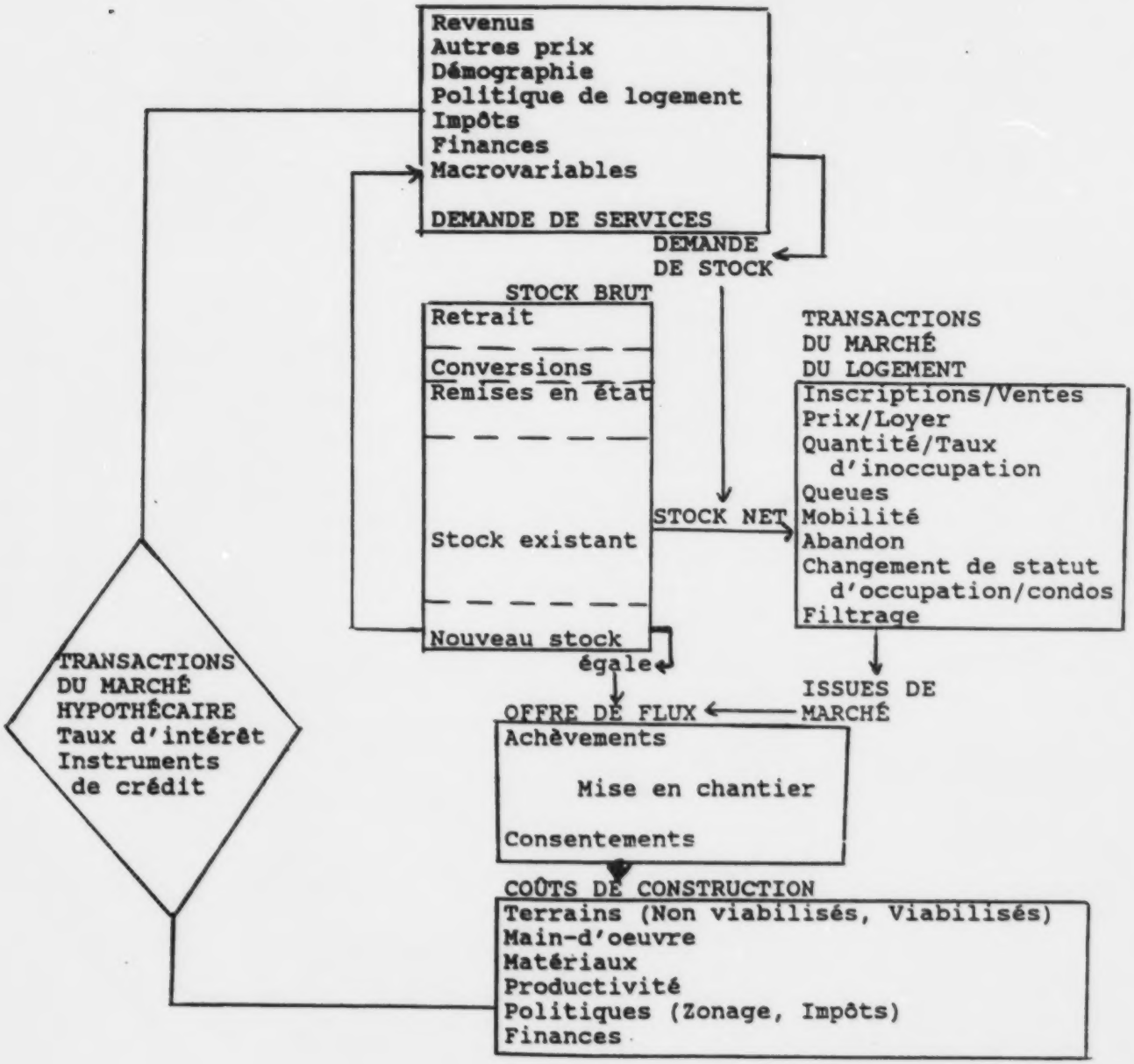
Notre examen de la disponibilité des données révèle que malgré l'existence de certains domaines où les données sont satisfaisantes, de nombreuses variables laissent sérieusement à désirer quant à la disponibilité et/ou la fiabilité des données à des fins de modélisation économétrique. Cela est surtout vrai des données qui sont disponibles annuellement. Les sources de données moins fréquentes (Recensement, EDF) sont plus fiables, mais ne peuvent servir qu'à vérifier l'exactitude des autres sources.

Des données sur de nombreuses variables domiciliaires et hypothécaires clés sont disponibles sur une base trimestrielle ou mensuelle et devraient pouvoir fournir une assise pour la construction d'un modèle économétrique de petite échelle du secteur du logement et du marché hypothécaire. La base de données disponibles est manifestement inadéquate pour l'élaboration d'un modèle économétrique trimestriel global. On ne peut se procurer que sur une base annuelle des données sur de nombreuses séries importantes, notamment le stock de logements et plusieurs variables qui agissent sur lui : retraits, dépréciation, ajouts et altérations. Même les données annuelles disponibles sont de mauvaise qualité. L'état des choses en matière de disponibilité des données limite certainement les perspectives de modélisation du secteur du logement et du marché hypothécaire.

La qualité des prévisions dépend très étroitement de la qualité et de l'étendue des données utilisées pour la modélisation. Nous recommandons vivement à la SCHL qu'elle crée et gère une banque de données complètes sur les variables du marché domiciliaire et hypothécaire. Certaines séries devront peut-être être élaborées ou estimées à partir de plus d'une source. Pour d'autres telles que les ménages, la SCHL devra peut-être commencer à rassembler elle-même les données nécessaires.

Figure 2.1

CADRE CONCEPTUEL DU MARCHÉ DU LOGEMENT
À DES FINS PRÉVISIONNELLES



3. PRÉVOIR AVEC UN MODÈLE UNIVARIÉ

Une approche simple et flexible de la prévision consiste à prédire les valeurs futures d'une variable quelconque en se fondant sur ses valeurs passées. On ajuste tout d'abord un modèle à série temporelle aux données et l'on extrapole ensuite la série vers l'avenir au moyen du modèle ajusté. Un tel modèle est « à base de données » en ce sens qu'il est spécifié à partir de la qualité de l'ajustement aux séries de données. Contrairement à ce qui est le cas en modélisation économétrique, comme nous l'avons vu au chapitre précédent, la théorie économique ne joue ici qu'un rôle modeste pour la spécification du modèle.

Un des nombreux types de modèles à séries temporelles fournissant des prévisions par extrapolation, à savoir le modèle autorégressif moyenne mobile intégré univarié (modèle ARIMA), connu également sous le nom de modèle « Box-Jenkins », s'est avéré en pratique très utile[2]. Dans le présent chapitre, nous (1) décrivons un modèle ARIMA des mises en chantier au Canada que nous avons élaboré et (2) analysons les avantages et inconvénients du modèle ARIMA comme outil prévisionnel. On trouvera dans de nombreux ouvrages un examen approfondi des modèles ARIMA, notamment ceux de Box et Jenkins (1970), Vandaele (1983), Granger et Newbold (1986) et Park 1989).

3.1 UN MODÈLE ARIMA DES MISES EN CHANTIER

A. Modélisation saisonnière multiplicative ARIMA

Le modèle autorégressif moyenne mobile (modèle ARMA) d'une série temporelle associe la valeur présente d'une variable à ses valeurs passées et aux valeurs présentes et passées d'erreurs aléatoires. On écrit souvent comme suit le processus ARMA d'ordre (p,q) :

$$Z_t = \phi_1 Z_{t-1} + \ldots + \phi_p Z_{t-p} + u_t - \theta_1 u_{t-1} - \ldots \theta_q u_{t-q}, \qquad (3.1a)$$

où Z_t, la valeur à l'instant t d'une série temporelle, est fonction de ses valeurs passées $Z_{t-1}, \ldots, Z_{t-p}$, et des erreurs aléatoires présentes et passées $u_t, u_{t-1}, \ldots, u_{t-q}$. Les θ_i et les θ représentent respectivement les paramètres autorégressifs et de

---

[2] On trouve dans la littérature un grand nombre de procédures prévisionnelles extrapolatives. Makridakis et al. (1982) font état de 24 méthodes extrapolatives à série temporelle utilisées dans un concours de prévision.

moyenne mobile. Les u_t sont des termes d'erreur aléatoires sériellement non corrélés de moyenne nulle et de variance constante σ^2, termes qu'on nomme « bruit blanc » en analyse de série temporelle.

Pour simplifier la notation, nous introduisons un opérateur de retard B dans l'indice intérieur temporel, de sorte que $BZ_t = Z_{t-1}$. Nous pouvons donc récrire (3.1a)

$$Z_t - \phi_1 B Z_t - \ldots - \phi_p B^p Z_t = a_t - \theta_1 B u_t - \ldots - \theta_q B^q u_t \quad (3.1b)$$

où, de manière plus concise,

$$\phi(B) Z_t = \theta(B) u_t, \quad (3.1c)$$

où $\phi(B)$ et $\theta(B)$ sont des polynômes en B d'ordre p et q, respectivement :

$$(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$$

et

$$\theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q.$$

On suppose que le processus ARMA (p,q) en (3.1) est « (faiblement) stationnaire » en ce sens que la moyenne, la variance et les autocorrélations ne changent pas dans le temps. La conséquence pratique de cette propriété est que la moyenne, la variance et les autocorrélations du processus peuvent être estimées avec cohérence à partir d'une série unique d'observations.

Il est courant d'observer dans les séries temporelles économiques comme celles des mises en chantier une tendance plutôt graduelle indiquant un changement lent de la moyenne dans le temps. Il est souvent possible de transformer une telle série non stationnaire en série stationnaire en la différenciant d fois.

$$Z_t = (1 - B)^d X_t$$

où d est un nombre entier non négatif. Si le processus ARMA (p,q) pour Z_t en (3.1) a été obtenu en différenciant la série originelle en X_t d fois, X_t est obtenue en sommant les séries en Z_t d fois : on parle dans ce cas de processus (ARIMA) autorégressif moyenne mobile intégré d'ordre (p, d, q).

En utilisant la notation pour un modèle ARMA en (3.1), nous pouvons définir un modèle ARIMA (p, d, q) comme suit :

$$(1 - \phi_1 B - \ldots - \phi_p B^p)(1 - B)^d X_t = (1 - \theta_1 B - \ldots - \theta_q B^q) u_t \quad (3.2)$$

où les u_t désignent un processus de type bruit blanc à moyenne nulle et à variance σ^2. On peut étendre légèrement (3.2) en y introduisant un terme constant :

$$(1-\phi_1 B-\ldots-\phi_p B^p)(1-B)^d X_t = \delta+(1-\theta_1 B-\ldots-\theta_q B^q)u_t \qquad (3.3)$$

Les séries temporelles économiques sont souvent caractérisées par une tendance exponentielle avec accroissement graduel de la dispersion dans le temps. Dans ce cas la transformation logarithmique de la série originelle pourra déboucher sur la stationnarité quant à la variance.

Une autre composante courante des séries temporelles économiques est une composante nommée « saisonnalité ». Il s'agit de toute variation cyclique ou périodique dans une série temporelle qui se répète selon une période fixe. De nombreuses séries économiques mensuelles ou trimestrielles sont caractérisées par des composantes saisonnières marquées. La série des mises en chantier possède une période saisonnière de douze mois. Les modèles ARIMA multiplicatifs sont utiles pour modéliser les séries temporelles à variations saisonnières de période connue s.

Si la série mensuelle ou trimestrielle considérée possède une périodicité annuelle, il sera peut-être nécessaire de procéder à une différenciation saisonnière pour obtenir la stationnarité de la moyenne. Il pourra s'avérer nécessaire d'appliquer la différenciation saisonnière

$$X_t - X_{t-s} = (1-B^s)X_t$$

D fois pour supprimer la tendance saisonnière de la série. Il pourra en outre être nécessaire d'avoir recours à une différenciation régulière ou non saisonnière d'ordre d, (1 - B) de la série différenciée saisonnièrement pour obtenir la stationnarité de la moyenne. Le modèle autorégressif moyenne mobile multiplicatif saisonnier intégré d'ordre p, d, q, P, D, Q, qu'on désigne ARIMA (p, d, q) x (P, D, Q), s'écrit comme suit :

$$\phi(B)\Phi(B)(1-B)^d(1-B^s)^D X_t = \theta(B)\Theta(B^s)u_t \qquad (3,4)$$

où $\Phi(B^s)$ et $\Theta(B^s)$ sont des polynômes en B^s de degré P et Q, respectivement, défini par

$$\Phi(B^s) = 1-\Phi_1 B^s-\ldots-\Phi_p B^{Ps}$$

et

$$\Theta(B^s) = 1-\Theta_1 B^s-\ldots-\Theta_q B^{Qs}$$

et où les polynômes $\phi(B)$ et $\theta(B)$ sont tels que défini en (3.1). Le modèle ARIMA en (3.4) et du même type que celui que nous avons élaboré pour les mises en chantier au Canada[3].

La série temporelle tracée à la Figure 3.1 donne le nombre mensuel de mises en chantier dans les centres urbains canadiens de 10 000 habitants ou plus (série CANSIM D849795) de janvier 1965 à décembre 1989. Les données sont exprimées en nombre total d'unités et ne sont pas ajustées pour tenir compte des variations saisonnières. Une inspection visuelle du tracé de la série temporelle indique une lente dérive de la moyenne et de la variance de même qu'une très forte composante saisonnière représentée par des pics et des vallées à des intervalles qui sont des multiples de 12 mois. Nous avons calculé les logarithmes communs de la série d'origine pour obtenir la stationnarité de la variance. Les séries logarithmiques sont figurées à la Figure 3.1(b).

À l'étape de l'identification d'un modèle ARIMA nous avons utilisé divers outils d'analyse des données pour parvenir à de premières estimations de la transformation des données, du nombre de dérivations pour obtenir la stationnarité et de l'ordre des polynômes d'autorégression et de mobilité de la moyenne du modèle (respectivement les polynômes AR et MA). Les instruments d'analyse des données à des fins d'identification comprennent les tracés des séries temporelles, les fonctions d'autocorrélation de l'échantillon, les fonctions d'autocorrélation partielle de l'échantillon, les fonctions d'autocorrélation d'un échantillon étendu (Tsay et Tiao (1984)) et les autocorrélations inverses (Cleveland (1972)).

Nous avons constaté qu'une procédure de sélection de modèle au moyen de ces instruments est généralement efficace, surtout lorsque complétée par des critères de sélection de modèle tels que le critère d'erreur de prédiction finale (FPE) d'Akaike (1969, 1970) et le critère d'information (AIC) d'Akaike (1974). Mais puisque le modèle ARIMA est spécifié sur la base des seules données, le concepteur de modèle doit souvent avoir recours à son jugement. Shibata (1985) a brièvement passé en revue diverses techniques de sélection de modèle en analyse de séries temporelles.

---

[3] Parmi les premiers exemples de modélisation ARIMA univariée appliquée au marché du logement figurent les textes classiques d'Abraham et Ledolter (1983), Pankratz (1983) et Vandaele (1983). Plus récemment, Sklarz et al. (1987) et Puri et Van Lierop (1988) ont élaboré des modèles ARIMA pour les séries mensuelles américaines des mises en chantier et ont constaté que leur efficacité prévisionnelle était satisfaisante.

Nous avons choisi d'ajuster un modèle ARIMA aux 288 observations de la période s'étendant de janvier 1965 à décembre 1988, en réservant 1989 comme période post-échantillonnage servant à vérifier la capacité prévisionnelle du modèle ajusté. Les fonctions d'autocorrélation et d'autocorrélation partielle (FAC et FACP) de l'échantillon des séries transformées tracées aux Figures 3.2(a) et (b) indiquent la non-stationnarité de la moyenne. Une autocorrélation partielle très marquée et significative de retard 1 et des autocorrélations diminuant lentement avec des retards saisonniers de 12, 24 et 36 suggèrent une différenciation unique de la série transformée à la fois régulièrement et de manière saisonnière. Cela a réduit le nombre total d'observations à T = 275.

Les FAC et FACP de l'échantillon des séries dérivées de la Figure 3.3 suggèrent un modèle ARIMA multiplicatif (1, 1, 1) x $(0, 1, 1)_{1,2}$ ou ARIMA (1, 1, 1) x $(3, 1, 0)_{1,2}$. Nous désignons ces modèles provisoires Modèle 1 :

$$(1-\phi_1 B)(1-B)(1-B^{12})\,LOGHS_t = (1-\theta_1 B)(1-\Theta_1 B^{12})\,u_t$$

et Modèle 2 :

$$(1-\phi_1 B)(1-\Phi_1 B^{12}-\Phi_2 B^{24}-\Phi_3 B^{36})(1-B)(1-B^{12})\,LOGHS_t = (1-\theta_1 B)(1-\Theta_1 B^{12})\,u_t$$

où $LOGHS_t$ est le logarithme commun des mises en chantier d'origine à l'instant t.

Une fois qu'un modèle ARIMA(p, d, q) x $(P, D, Q)_s$ est identifié, on peut estimer ses paramètres au moyen de la méthode des moindres carrés (MC) conditionnels. En supposant la normalité des erreurs aléatoires, on peut également obtenir des estimations approximatives de vraisemblance maximale. Ce caractère approximatif est la résultant des suppositions qu'il faut faire à propos des valeurs initiales des X_t et des u_t. Voir Box et Jenkins (1970). On peut calculer des estimations exactes de la vraisemblance maximale (VM) de diverses manières, proposées notamment par Newbold (1974), Dent (1977) et Ansley (1979). Quoique ces estimateurs soient asymptotiquement équivalents en ce sens qu'ils ont les mêmes distributions de probabilité dans les grands échantillons, leurs estimations peuvent varier considérablement entre elles.

Les estimations de VM pour le Modèle provisoire 1 sont résumées au Tableau 3.1(A)[4]. L'estimation du coefficient

---

[4] Les calculs dont il est fait état aux Chapitres 3 et 4 ont été fournis par le programme SCA-UTS de Scientific Computing Associates. Leur méthode d'estimation est fondée sur une approche de vraisemblance maximale conditionnelle. Voir Liu et al. (1986).

d'autorégression est significative et satisfait la condition de stationnarité. En revanche, les deux coefficients MA sont significatifs et satisfont les conditions d'« inversibilité », bien que $\hat{\Theta}_1 = .9192$ soit proche de la limite de non-inversibilité[5].

Aucun modèle n'est « fidèle ». Aussi l'efficacité des modèles doit-elle être soigneusement vérifiée. Parmi les instruments de vérification diagnostique figurent les autocorrélations résiduelles et les autocorrélations partielles, le test portemanteau de Box-Pierce-Ljung pour défaut d'ajustement et les tests au moyen du multiplicateur de Lagrange en cas de diverses autres éventualités spécifiques. Newbold (1983b) a passé en revue les tests diagnostiques utilisés en modélisation de séries temporelles univariées ou multivariées. Quoique le test de Box-Pierce-Ljung cadrât avec l'hypothèse que la série de termes résiduels du modèle ajusté constitue un processus semblable au bruit blanc, les FAC et les FACP comportaient un coefficient significatif au retard 36 et laissent supposer que le modèle 2 apporte peut-être une meilleure spécification.

Les estimations des paramètres pour le Modèle provisoire 2 sont les suivantes :

$\hat{\phi}_1$ = 0,6996 où $\underline{t}$ = 8,60
$\hat{\Phi}_1$ = -,1621 où $\underline{t}$ = -2,04
$\hat{\Phi}_2$ = -,0661 où $\underline{t}$ = -0,90
$\hat{\Phi}_3$ = -,2503 où $\underline{t}$ = 3,85
$\hat{\theta}_1$ = 0,9003 où $\underline{t}$ = 18,19
$\hat{\Theta}_1$ = 0,6850 où $\underline{t}$ = 9,82

et avec SEE = 0,0658. Nous constatons que l'estimation du paramètre autorégressif saisonnier (2) n'est pas significative pour $\alpha = 0.05$.

Nous avons procédé à une vérification supplémentaire de l'adéquation du modèle en imposant diverses contraintes. En estimant un modèle ARIMA(1, 1, 1) x $(3, 1, 1)_{12}$ sans constante et avec une contrainte $\Phi_2 = 0$, nous avons observé que l'estimation du paramètre SAR(1) est statistiquement non significative. Nous

---

[5] On dit qu'un modèle ARMA est inversible si la série de coefficients dans sa représentation autorégressive pure converge lorsque le temps écoulé approche de l'infini. Dans le Modèle 1, les conditions d'inversibilité sont $-1 < \theta < 1$ et $-1 < \Theta < 1$.

conformant au principe d'économie, nous avons choisi le modèle ARIMA (1, 1, 1) x $(3, 1, 1)_{12}$ avec pour contraintes $\Phi_1 = \Phi_2 = 0$:

$$(1-\phi_1 B)(1-\Phi_3 B^{36})(1-B)(1-B^{12})LOGHS_t = (1-\theta_1 B)(1-\Theta_1 B^{12})u_t \quad (3.5)$$

Nous avons présenté au Tableau 3.1 (B) les résultats de l'estimation du modèle en (3.5). Toutes les estimations de coefficients sont significatives, elles satisfont toutes les conditions de stationnarité et d'inversibilité et l'erreur-type résiduelle est réduite de SEE = ,0676 à SEE = ,0656. Le modèle ajusté devient alors

$$(1-0.7232B)(1+0.2119B^{36})(1-B)(1-B^{12})LOGHS_t = (1-0.9174B)(1-0.7773B^{12})u_t \quad (3.6)$$

On trouvera à la Figure 3.4 les FAC et FACP de l'échantillon pour les résidus du modèle estimé (3.6). Les autocorrélations et les autocorrélations partielles restent significatives au retard 36. Toutefois, les statistiques de Box-Pierce-Ljung ont pour valeur Q(12) = 4,9,Q(24) = 12,2 et Q(36) - 25,2 et confortent sérieusement l'hypothèse que les résidus se comportent comme du bruit blanc. Le modèle estimé en (3.6) semble adéquat.

On note que l'estimation du paramètre MA(1) non saisonnier est proche de la limite du domaine de non-inversibilité. En général, de nombreuses séries temporelles économiques ne tendent à devenir stationnaires qu'après différenciation, or la différenciation engendre souvent une composante MA pour laquelle les paramètres sont situés très près de la limite du domaine de non-inversibilité. Un modèle ARIMA dans lequel certaines des racines du polynôme MA sont égales à 1 est parfaitement valide, et il n'est peut-être pas raisonnable d'exiger que les estimations des paramètres MA satisfassent aux conditions d'inversibilité.

3.2 PRÉVOIR AVEC UN MODÈLE ARIMA

Une fois élaboré un modèle ARIMA adéquat, nous pouvons en tirer des prévisions de manière mécanique. La marche à suivre pour obtenir des prévisions est décrite ci-dessous à propos du modèle ARMA de base, mais les mêmes principes s'appliquent lorsqu'on a affaire au modèle ARIMA saisonnier multiplicatif.

Supposons pour l'instant que les paramètres du modèle soient connus et que nous ayons besoin à l'instant T d'une prévision pour h périodes vers l'avenir d'un processus ARMA Z_t. Si le modèle est valide dans le futur, alors

$$Z_{T+h} = \phi_1 Z_{T+h-1} + \ldots + \phi_p Z_{T+h-p} + u_{T+h} - \theta_1 u_{T+h-1} - \ldots - \theta_q u_{T+h-q} \quad (3.7)$$

À l'instant de la prévision T, nous pourrons avoir un ensemble d'informations disponibles appelé « ensemble d'informations » que nous pourrons utiliser pour prévoir Z_{T+h}. Par exemple, l'ensemble d'informations, noté I_T, pourra consister dans tout l'historique de la série jusqu'à l'instant T : Z_T, Z_{T-1}, ...

Si le processus de bruit blanc u_t est normal, la prévision « optimale » (en ce sens que l'erreur en moindres carrés est minimale) à l'instant T de Z_{T+h} dépendante de l'ensemble d'information I_T est tout simplement son espérance conditionnelle

$$f_T(h) = E_T[Z_{T+h} | I_T] \tag{3.8}$$

où $f^T(h)$ représente la prévision à h périodes vers l'avenir à l'instant T et $E_T[.]$ désigne l'espérance conditionnelle mesurée à l'instant T. En reprenant l'espérance conditionnelle de (3.7), nous pouvons écrire :

$$f_T(h) = E_T(Z_{T+h}) = \phi_1 E_T(Z_{T+h-1}) + \ldots + \phi_p E_T(Z_{T+h-p}) \tag{3.9}$$
$$+ E_T(u_{T+h}) - \theta_1 E_T(u_{T+h-1}) - \ldots - \theta_q E_T(u_{T+h-q})$$

L'espérance conditionnelle de la future Z_t étant donné I_T est sa valeur prévisionnelle, alors que celle, présente ou passée, de Z_t est égale à sa valeur effective.

Il est clair que l'erreur prévisionnelle à h étapes vers l'avenir, $e_T(h)$, est donné par :

$$e = Z_{T+h} - f_T(h)$$
$$= (1 + \psi_1 + \ldots + \psi_{h-1} B^{h-1}) u_{T+h}, \tag{3.10}$$

où les ψ sont les coefficients obtenus lorsqu'on exprime le processus Z_t en termes des erreurs aléatoires présentes et passées uniquement. L'erreur prévisionnelle à h étapes vers l'avenir a donc pour moyenne la valeur zéro et pour variance

$$Var[e_T(h)] = (1 + \psi_1^2 + \ldots + \psi_{h-1}^2)\sigma^2. \tag{3.11}$$

Contrairement à ce qui est le cas pour les prévisions fondées sur un modèle économétrique, nous pouvons facilement calculer les erreurs-types des prévisions ARIMA, de sorte qu'une mesure de précision est associée aux prévisions ponctuelles. Si le processus u_t est normal, alors $e_T(h)$ est normale avec une moyenne égale à zéro et une variance donnée par (3.11) et fournit une base pour les intervalles prévisionnels. Un intervalle prévisionnel de 95 p. 100 s'écrira par exemple

$$f_T(h) \pm 1.96\left[\left(1+\psi_1^2+\ldots+\psi_{h-1}^2\right)\sigma^2\right]^{1/2} \qquad (3.12)$$

Les considérations précédentes sur les prévisions s'appuient sur la supposition que tout l'historique du processus ARIMA est connu. Mais en pratique ce n'est pas le cas. Le fait que l'ensemble de données disponibles ne consiste qu'en T observations, X_1, ..., X_T, pose deux problèmes. Premièrement, les erreurs prévisionnelles u_1, u_2, ..., u_T ne sont plus à notre disposition et doivent donc être estimées. Si nous élaborons la séquence de prévisions à partir du début de la série de données et si le processus ARIMA considéré est inversible, les erreurs prévisionnelles engendrées par les séries u_t estimées devraient être négligeables. Deuxièmement, puisque l'historique entier de la série n'est pas à notre disposition, les paramètres du modèle sont inconnus et doivent être estimés. Il faut donc estimer également les espérances conditionnelles et la variance prévisionnelle. Des ambiguïtés dans la spécification du modèle et des erreurs d'estimation biaisent les prévisions et entraînent la sous-estimation de la valeur réelle de la variance prévisionnelle. En général, toutefois, l'identification et l'estimation d'un modèle ARIMA deviennent de plus en plus précises à mesure que s'allonge la série temporelle à analyser.

On a utilisé le modèle ajusté (3.6) pour prévoir la série des mises en chantier en 1989. En résolvant (3.6) par multiplication et en réarrangeant les termes, nous pouvons écrire comme suit la fonction prévisionnelle à h étapes vers l'avenir à l'origine T

$$\begin{aligned} E_T(LOGHS_{T+h}) &= 1.7232E_T(LOGHS_{T+h-1}) - .7232E_T(LOGHS_{T+h-2}) \\ &+ E_T(LOGHS_{T+h-12}) - 1.7232E_T(LOGHS_{T+h-13}) \\ &+ .7232E_T(LOGHS_{T+h-14}) - .2119E_T(LOGHS_{T+h-36}) \\ &+ .3651E_T(LOGHS_{T+h-37}) - .1532E_T(LOGHS_{T+h-38}) \\ &+ .2119E_T(LOGHS_{T+h-48}) - .3651E_T(LOGHS_{T+h-49}) \\ &+ .1532E_T(LOGHS_{T+h-50}) + E_T(u_{T+h}) \\ &- .9172E_T(u_{T+h-1}) - .7773E_T(u_{T+h-12}) \\ &+ .7129E_T(u_{T+h-13}) \end{aligned} \qquad (3.13)$$

en termes d'espérances conditionnelles. Si nous voulons obtenir les prévisions dans leurs unités d'origine plutôt que leurs valeurs logarithmiques, il suffit de déterminer l'antilogarithme des prévisions des séries $LOGHS_t$.

On trouvera au Tableau 3.2 un profil prévisionnel sur douze mois de même qu'un ensemble d'intervalles prévisionnels de confiance de 95 p. 100 pour les douze mois de 1989, l'origine étant en décembre 1988. Étant donné que ces prévisions ont été engendrées par le modèle en (3.6) fondé sur les données pour 1965-1988, ce sont des prévisions strictement post-échantillonnage. La Figure 3.5 illustre les tracés de ces

prévisions. Il est remarquable que malgré tous les changements et fluctuations dans les mises en chantier en 1989, le tracé prévisionnel sur douze mois reproduit très fidèlement les valeurs observées en 1989. De plus, la qualité des prévisions ne s'est pas du tout détériorée avec l'éloignement de l'horizon prévisionnel.

Les erreurs prévisionnelles effectives varient en valeur absolue de seulement 0,081 p. 100 pour la prévision 10 mois vers l'avenir d'octobre 1989 à 1,737 p. 100 pour la prévision un mois vers l'avenir de janvier 1989. Pour la période d'échantillonnage allant de janvier 1965 à décembre 1988, la déviation standard de la série transformée est de 0,1584 alors que celle de la série résiduelle est de 0,0656. Lorsque la série à prévoir contient une forte composante de bruit, nous estimons que les prévisions sont remarquablement précises. Nous concluons que le modèle ARIMA en (3.6) reproduit excellemment les données sur les mises en chantier au Canada.

Depuis 1974, Nelson (1984) réalise des prévisions sur trois variables macroéconomiques américaines, y compris le PNB, au moyen de modèles ARIMA, fournissant ainsi des prévisions repères pour les comparaisons avec les prévisions tirées de modèles économétriques à grande échelle. McNees (1988) constate que malgré que les prévisions de Nelson ne fussent aucunement ajustées subjectivement, leur efficacité est comparable à celle des prévisions ajustées subjectivement des prévisionnistes commerciaux éminents.

Il est largement admis que les modèles ARIMA simples sont aussi précis pour les prévisions à court terme que les grands modèles économétriques complexes fournissant des prévisions ajustées subjectivement. Nous avons en outre constaté que le modèle ARIMA que nous avons élaboré dans la présente étude prédit remarquablement bien les mises en chantier au Canada jusqu'à un horizon de douze mois.

Comme cela ressort clairement de l'exposé ci-dessus de la méthode de modélisation et de prévision ARIMA, cette méthode présente notamment comme outil prévisionnel les avantages suivants :

(1) elle fournit des prévisions très précises au moins à court terme;
(2) elle constitue un procédé remarquablement simple pour obtenir des prévisions;
(3) il est possible d'élaborer un modèle ARIMA en utilisant des données remontant à n'importe quel moment du passé et d'effectuer ensuite des prévisions permettant de

vérifier facilement la précision prévisionnelle du modèle;

(4) on peut cerner avec plus de précision la dynamique des variables;

(5) il est plus facile d'identifier les structures d'erreur et de les incorporer au modèle;

(6) le caractère saisonnier des séries temporelles est introduit dans le modèle de manière systématique; et

(7) on peut calculer pour chaque horizon prévisionnel les erreurs-types comme mesure de la précision prévisionnelle.

Il n'est donc pas étonnant que les modèles ARIMA prédisent remarquablement bien, tout au moins à court terme ou lorsque l'activité économique est relativement stable.

En revanche, la méthode de prévision ARIMA univariée présente les faiblesses suivantes :

(1) la précision prévisionnelle peut diminuer rapidement parallèlement à l'éloignement de l'horizon prévisionnel;

(2) les aspects non linéaires de l'économie peuvent dans certains cas lui échapper; et

(3) elle pourra ne pas être à la hauteur en période de changement structurel ou lorsque l'activité économique est instable à cause de chocs exogènes tels que guerres ou changements de politique inattendus.

La capacité prévisionnelle d'un modèle ARIMA ou de tout autre modèle à série temporelle se fonde sur la dépendance des observations successives dans la série. À mesure que s'éloigne l'horizon prévisionnel, le degré de corrélation entre la variable à prévoir et la série de données observées tend à diminuer, entraînant une détérioration de la capacité prévisionnelle du modèle ARIMA pour des horizons prévisionnels plus éloignés.

Les changements structurels subits en période de chocs exogènes tels que les conflits syndicaux et les embargos pétroliers constituent le type de non-stationnarité qui est le plus difficile à traiter en modélisation avec série temporelle. De tels changements peuvent induire un effet transitoire à court terme ou un changement à long terme dans la structure du modèle. Hilmer (1984) démontre comment un seul facteur perturbateur peut influer sur plusieurs prévisions consécutives tant qu'il n'aura pas été supprimé. La modélisation ARIMA en présence de tels facteurs perturbateurs a été étudiée par Change, Tiao et Chen (1988). Box et Tiao (1975) ont proposé une variante du modèle ARIMA, appelée « analyse interventionnelle », qu'on peut utiliser lorsque certains événements connus ont influé sur la série temporelle considérée.

Manifestement, les modèles ARIMA fournissent pour la prévision de séries chronologiques un cadre conceptuel très simple dans lequel la spécification d'un modèle n'est fondée que sur les seules données. Cette méthode est particulièrement utile lorsqu'il est difficile d'identifier les principaux facteurs qui influent sur la variable à prévoir et de modéliser leurs interrelations.

Tableau 3.1

(A) RÉSUMÉ D'UN MODÈLE UNIVARIÉ DES MISES EN CHANTIER
(En logarithmes vulgaires)

VARIABLE	TYPE DE VARIABLE	ORIGINEL OU CENTRE	DIFFÉRENCIATION
LOGHS	ALÉATOIRE	ORIGINEL	$(1-B)^{1}(1-B^{12})$

DÉSIGNATION DU PARAMÈTRE	NOM DE LA VARIABLE	VALEUR	ERREUR DÉVIATION STANDARD	VALEUR T
1 THETA1	LOGHS	0.9192	0.0370	24.83
2 THETA12	LOGHS	0.8462	0.0340	24.91
3 PHI1	LOGHS	0.7118	0.0670	10,63

SOMME TOTALE DES CARRÉS - - - - - - - - - - - - 0.731151E+01
NOMBRE TOTAL D'OBSERVATIONS - - - - - - - - - - 288
SOMME RÉSIDUELLE DES CARRÉS - - - - - - - - - - 0.124591E+01
R AU CARRÉ - - - - - - - - - - - - - - - - - - 0,821
NOMBRE EFFECTIF D'OBSERVATIONS - - - - - - - - 274
ESTIMATION DE LA VARIANCE RÉSIDUELLE - - - - - - 0.454712E-02
ERREUR-TYPE RÉSIDUELLE - - - - - - - - - - - - 0,674323E-01

(B) RÉSUMÉ D'UN MODÈLE UNIVARIÉ DES MISES EN CHANTIER
(En logarithmes vulgaires)

VARIABLE	TYPE DE VARIABLE	ORIGINEL OU CENTRE	DIFFÉRENCIATION
LOGHS	ALÉATOIRE	ORIGINEL	$(1-B)^{1}(1-B^{12})$

DÉSIGNATION DU PARAMÈTRE	NOM DE LA VARIABLE	VALEUR	ERREUR DÉVIATION STANDARD	VALEUR T
1 THETA1	LOGHS	0.9174	0.0421	21.78
2 THETA12	LOGHS	0.7773	0.0435	17.86
3 PHI1	LOGHS	0.7232	0.0734	9.86
4 PHI36		-0.2119	0.0621	-3.42

SOMME TOTALE DES CARRÉS 0.731E+01 NOMBRE EFFECTIF D'OBSERVATIONS 238
ESTIMATION DE LA VARIANCE RÉSIDUELLE 0.43E-2 SOMME RÉSIDUELLE DES CARRÉS 0.102E+01
R AU CARRÉ 0,831 NOMBRE TOTAL D'OBSERVATIONS 288
ERREUR-TYPE STANDARD 0.65476E-01

Tableau 3.2

PROFIL PRÉCISIONNEL AVEC HORIZON DE 12 MOIS
(En logarithmes vulgaires)

MOIS 1989	LIMITE INFÉRIEURE POURCENTAGE	VALEUR DE PRÉVISION	LIMITE SUPÉRIEURE	VALEUR EFFECTIVE	ERREUR DE PRÉVISION	ERREUR DE PRÉVISION
Jan.	3.936	4.064	4.192	4.136	.072	1.737
Fév.	3.786	3.951	4.116	3.990	.039	.982
Mar.	3.846	4.032	4.218	4.085	.053	1.290
Avr.	3.997	4.196	4.395	4.240	.044	1.046
Mai	4.094	4.303	4.512	4.282	-.021	-.489
Juin	4.057	4.273	4.489	4.240	-.033	-.779
Juil.	4.000	4.222	4.444	4.215	-.007	-.156
Août	3.998	4.225	4.453	4.190	-.036	-.850
Sep.	3.959	4.191	4.423	4.213	.022	.531
Oct.	3.959	4.194	4.430	4.198	.003	.081
Nov.	3.982	4.221	4.461	4.193	-.028	-.662
Déc.	3.918	4.161	4.404	4.140	-.021	-.515

Note : Les prévisions s'appliquent aux 12 mois de l'année 1989; elles sont fondées sur des données dont les plus récentes datent de décembre 1988.

Figure 3.1 (A) Mises en chantier (nombre)

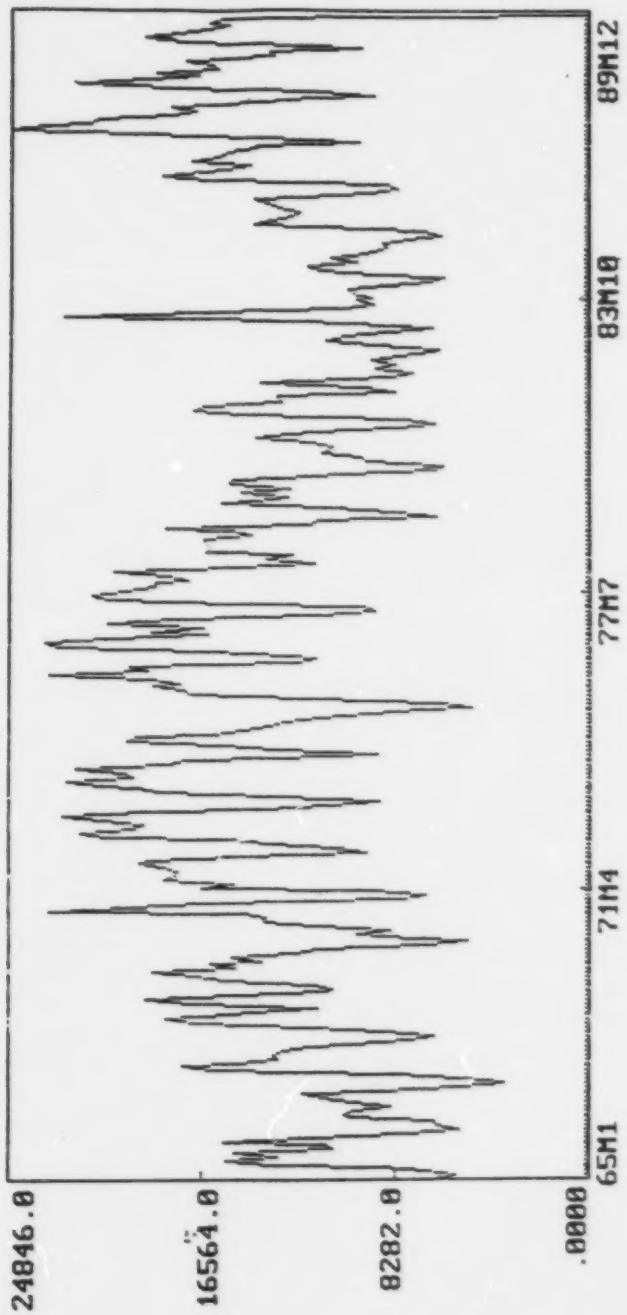


Figure 3.1 (B) Mises en chantier (logarithme)

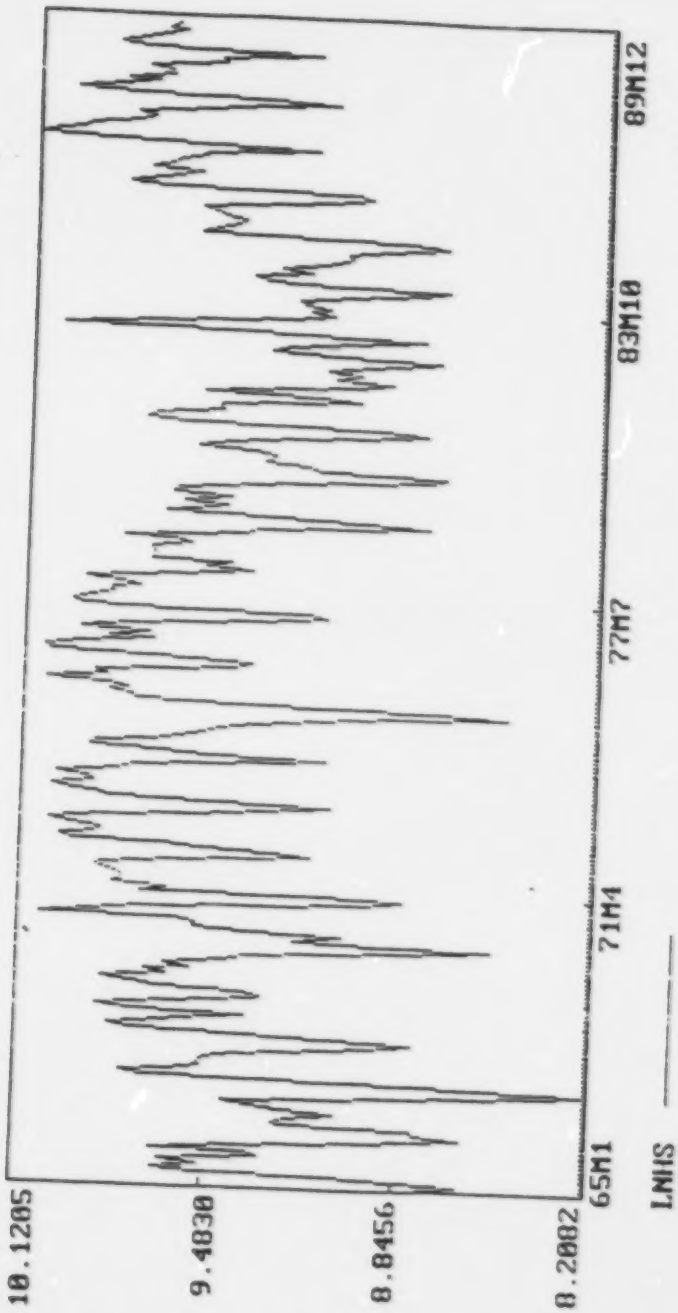


Figure 3.1 (A)

AUTOCORRÉLATIONS D'ÉCHANTILLON DES MISES EN CHANTIER
(En logarithmes vulgaires)

PÉRIODE DE TEMPS ANALYSÉE - - - - - - - - - - - - 1 to 288
NOM DE LA SÉRIE - - - - - - - - - - - - - - - - - LOGHS
NOMBRE EFFECTIF D'OBSERVATIONS - - - - - - - - - 288
DÉVIATION STANDARD DE LA SÉRIE - - - - - - - - - 0,1593
MOYENNE DE LA SÉRIE (DIFFÉRENCIÉE) - - - - - - - 4,1014
DÉVIATION STANDARD DE LA MOYENNE - - - - - - - - 0.0094
VALEUR T DE LA MOYENNE (PAR RAPPORT À ZÉRO) - - 436.8426

1-12	.79	.47	.22	.13	.14	.15	.11	.06	.09	.28	.52	.62
Er.-T.	.06	.09	.10	.10	.10	.10	.10	.10	.10	.10	.10	.11
Q	181	245	259	264	270	276	280	281	283	307	389	504
13-24	.47	.22	.02	-.03	.01	.03	-.00	-.03	.01	.20	.44	.54
Er.-T.	.12	.13	.13	.13	.13	.13	.13	.13	.13	.13	.13	.14
Q	572	587	587	587	587	587	587	588	588	601	662	753
25-36	.40	.13	-.07	-.12	-.08	-.07	-.11	-.16	-.12	.05	.27	.36
Er.-T.	.14	.15	.15	.15	.15	.15	.15	.15	.15	.15	.15	.15
Q	803	808	810	815	817	819	823	831	835	836	861	904

```
          -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
           +----+----+----+----+----+----+----+----+----+----+----+
 1   .79                          +  IXX+XXXXXXXXXXXXXXXXX
 2   .47                         +   IXXX+XXXXXXXX
 3   .22                        +    IXXXXX
 4   .13                        +    IXXX +
 5   .14                        +    IXXX +
 6   .15                        +    IXXXX+
 7   .11                        +    IXXX +
 8   .06                        +    IX   +
 9   .09                        +    IXX  +
10   .28                        +    IXXXX+XX
11   .52                        +    IXXXX+XXXXXXXX
12   .62                       +     IXXXXX+XXXXXXXXX
13   .47                       +     IXXXXX+XXXXXX
14   .22                       +     IXXXXX+
15   .02                       +     IX    +
16  -.03                       +    XI     +
17   .01                       +     I     +
18   .03                       +     IX    +
19   .00                       +     I     +
20  -.03                       +    XI     +
21   .01                       +     I     +
22   .20                       +     IXXXXX+
23   .44                       +     IXXXXX+XXXXX
24   .54                      +      IXXXXXX+XXXXXX
```

Figure 3.2 (B)

AUTOCORRÉLATIONS PARTIELLES D'ÉCHANTILLON
DES MISES EN CHANTIER
(En logarithmes vulgaires)

1-12	.79	-.41	.05	.18	.01	-.03	-.09	.10	.21	.38	.21	-.05
Er.-T.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
13-24	-.28	-.10	.01	.04	-.01	-.07	.06	.05	.01	.13	.17	-.02
Er.-T.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
25-36	-.17	-.16	.00	.01	-.02	-.08	-.03	.01	-.04	-.04	.12	-.02
Er.-T.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

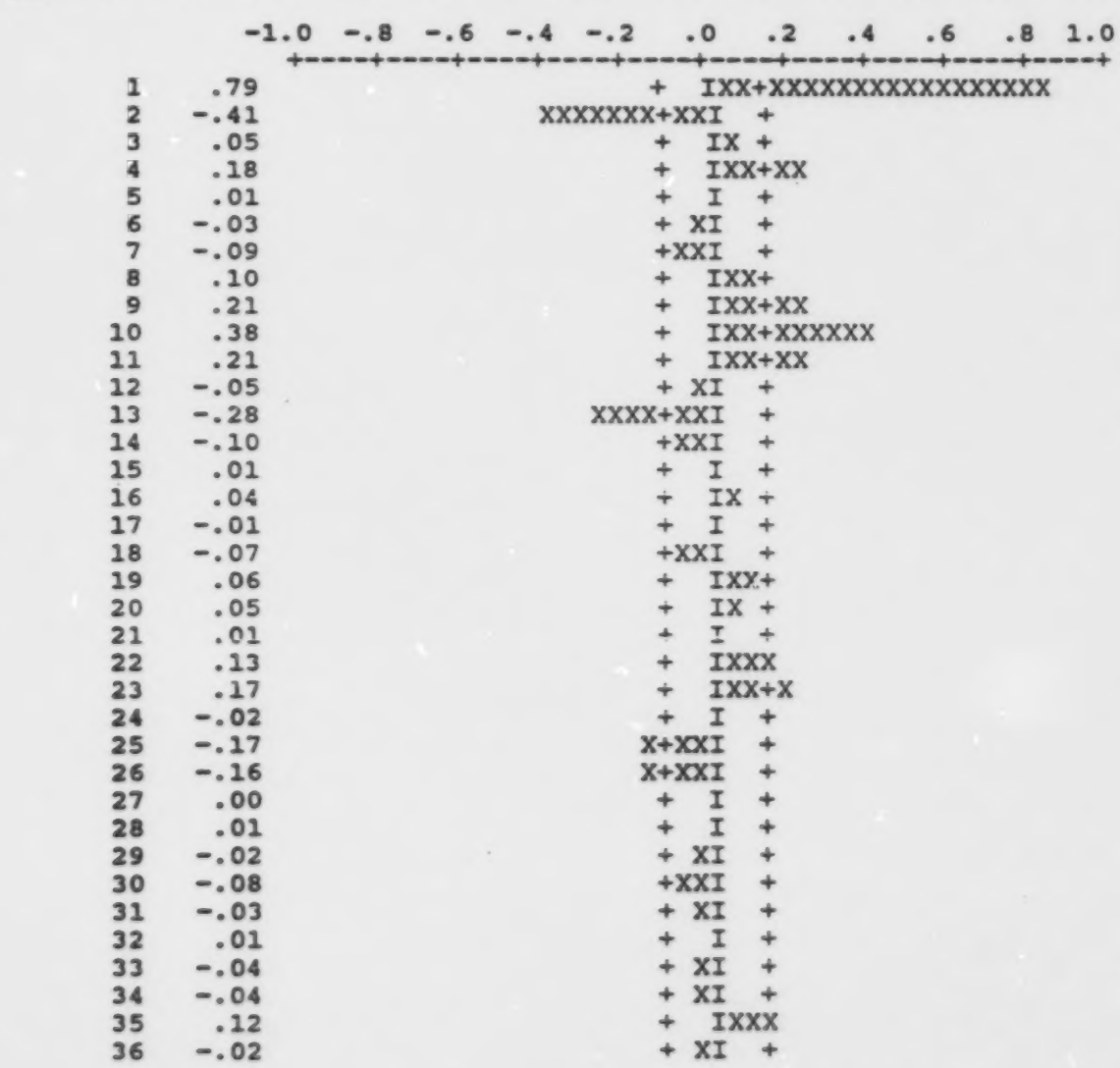


Figure 3.3 (A)

AUTOCORRÉLATIONS D'ÉCHANTILLON DES MISES EN CHANTIER TRANSFORMÉES
(En logarithmes vulgaires, d = D = 1)

ORDRES DE DIFFÉRENCIATION . . . . . $(1-B^{1})(1-B^{12})$
PÉRIODE DE TEMPS ANALYSÉE . . . . . 1 to 288
NOM DE LA SÉRIE . . . . . LOGHS
NOMBRE EFFECTIF D'OBSERVATIONS 275
DÉVIATION STANDARD DE LA SÉRIE 0.0916
MOYENNE DE LA SÉRIE (DIFFÉRENCIÉE) . . . . . -0.0003
DÉVIATION STANDARD DE LA MOYENNE 0.0055
VALEUR T DE LA MOYENNE (PAR RAPPORT À ZÉRO) . . . -0.0589

1-12	-.11	-.07	-.04	-.04	-.12	.03	.05	-.00	.00	.10	.05	-.50
Er.-T.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
Q	3.6	5.1	5.5	6.0	10.2	10.4	11.1	11.1	11.1	14.1	14.8	88.0
13-24	-.05	.04	.03	.01	.08	-.05	-.03	.07	.01	-.00	.00	.13
Er.-T.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08
Q	88.6	89.1	89.3	89.4	91.2	92.1	92.3	94.0	94.0	94.0	94.0	99.1
25-36	.04	-.02	-.07	-.01	-.00	.06	-.01	.01	-.02	-.12	.06	-.24
Er.-T.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08
Q	99.5	99.6	101	101	101	102	102	102	102	107	108	126

```
          -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
            +----+----+----+----+----+----+----+----+----+----+
 1   -.11                          XXXI  +
 2   -.07                          +XXI  +
 3   -.04                          + XI  +
 4   -.04                          + XI  +
 5   -.12                          XXXI  +
 6    .03                          +  IX +
 7    .05                          +  IX +
 8    .00                          +  I  +
 9    .00                          +  I  +
10    .10                          +  IXXX
11    .05                          +  IX +
12   -.50                XXXXXXXXXX+XXI  +
13   -.05                        +  XI   +
14    .04                        +   IX  +
15    .03                        +   IX  +
16    .01                        +   I   +
17    .08                        +   IXX +
18   -.05                        +  XI   +
19   -.03                        +  XI   +
20    .07                        +   IXX +
21    .01                        +   I   +
22    .00                        +   I   +
23    .00                        +   I   +
24    .13                        +   IXXX+
```

Figure 3.3 (B)

AUTOCORRÉLATIONS D'ÉCHANTILLON DES MISES EN CHANTIER TRANSFORMÉES
(En logarithmes vulgaires, d = D = 1)

1-12	-.11	-.09	-.06	-.06	-.15	-.02	.02	-.01	-.00	.09	.09	-.49
Er.-T.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
13-24	-.21	-.07	-.03	-.07	-.10	-.09	-.05	.05	-.01	.10	.10	-.12
Er.-T.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
25-36	-.09	-.04	-.08	-.07	-.02	-.03	-.08	.06	.03	-.09	.12	-.33
Er.-T.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

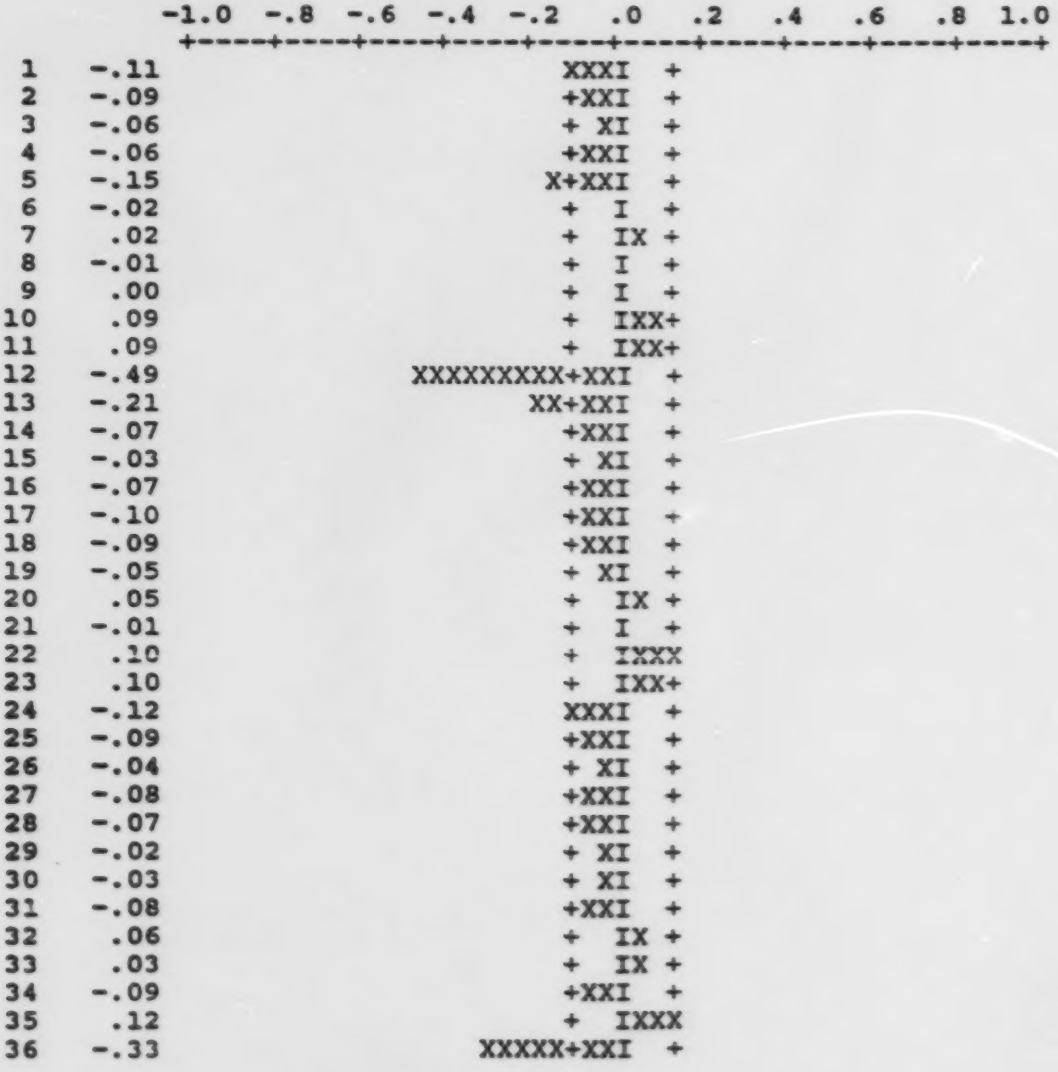


Figure 3.4 (A)

AUTOCORRÉLATIONS D'ÉCHANTILLON DES RÉSIDUS

PÉRIODE DE TEMPS ANALYSÉE . . . . . 51 To 288
NOM DE LA SÉRIE . . . . . RY 238
NOMBRE EFFECTIF D'OBSERVATIONS 0.0642
DÉVIATION STANDARD DE LA SÉRIE 0.0000
MOYENNE DE LA SÉRIE (DIFFÉRENCIÉE) 0.0042
DÉVIATION STANDARD DE LA MOYENNE . . . . . -0,0094
VALEUR T DE LA MOYENNE (PAR RAPPORT À ZÉRO) . .

1-12	-.01	.00	.04	.03	-.04	.04	.03	-.01	-.07	.04	.04	-.07
Er.-T.	06	.06	.06	.06	.06	.07	.07	.07	.07	.07	.07	.07
Q	.0	.0	.4	.6	.9	1.3	1.5	1.5	2.8	3.2	3.6	4.9
13-24	-.06	.04	.01	.01	-.01	-.09	-.08	.07	-.03	.01	.04	.03
Er.-T.	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07
Q	5.9	6.2	6.3	6.3	6.3	8.5	10.2	11.3	11.6	11.6	12.0	12.2
25-36	.07	-.02	-.05	.02	-.02	.03	-.05	.01	-.06	-.16	.05	-.03
Er.-T.	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07
Q	13.5	13.6	14.3	14.4	14.4	14.7	15.5	15.5	16.6	24.2	24.9	25.2

		-1.0 -.8 -.6 -.4 -.2 .0 .2 .4 .6 .8 1.0
		+----+----+----+----+----+----+----+----+----+----+----+
1	-.01	+ I +
2	.00	+ I +
3	.04	+ IX +
4	.03	+ IX +
5	-.04	+ XI +
6	.04	+ IX +
7	.03	+ IX +
8	-.01	+ I +
9	-.07	+XXI +
10	.04	+ IX +
11	.04	+ IX +
12	-.07	+XXI +
13	-.06	+XXI +
14	.04	+ IX +
15	.01	+ I +
16	.01	+ I +
17	-.01	+ I +
18	-.09	+XXI +
19	-.08	+XXI +
20	.07	+ IXX+
21	-.03	+ XI +
22	.01	+ I +
23	.04	+ IX +
24	.03	+ IX +

Figure 3.4 (B)

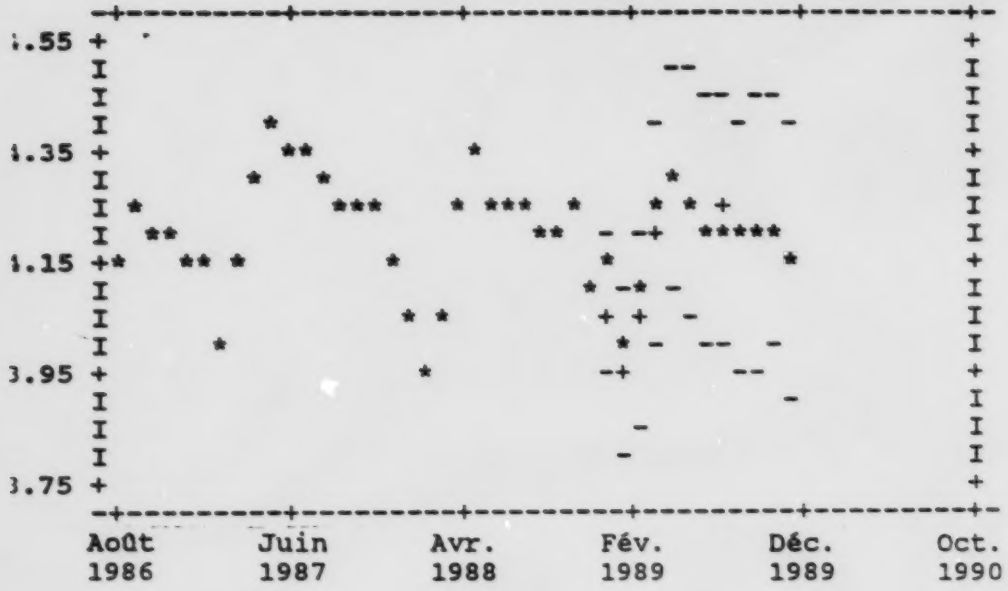
AUTOCORRÉLATIONS PARTIELLES D'ÉCHANTILLON DES RÉSIDUS

1-12	-.01	.00	.04	.03	-.04	.04	.02	-.01	-.07	.03	.04	-.07
Er.-T.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
13-24	-.07	.03	.03	.02	-.02	-.10	-.07	.08	-.04	.01	.05	.04
Er.-T.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
25-36	.08	-.03	-.07	.02	.01	.02	-.09	.02	-.04	-.17	.05	-.05
Er.-T.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

		-1.0 -.8 -.6 -.4 -.2 .0 .2 .4 .6 .8 1.0
		+----+----+----+----+----+----+----+----+----+----+----+
1	-.01	+ I +
2	.00	+ I +
3	.04	+ IX +
4	.03	+ IX +
5	-.04	+ XI +
6	.04	+ IX +
7	.02	+ IX +
8	-.01	+ I +
9	-.07	+XXI +
10	.03	+ IX +
11	.04	+ IX +
12	-.07	+XXI +
13	-.07	+XXI +
14	.03	+ IX +
15	.03	+ IX +
16	.02	+ IX +
17	-.02	+ I +
18	-.10	XXXI +
19	-.07	+XXI +
20	.08	+ IXX+
21	-.04	+ XI +
22	.01	+ I +
23	.05	+ IX +
24	.04	+ IX +
25	.08	+ IXX+
26	-.03	+ XI +
27	-.07	+XXI +
28	.02	+ IX +
29	.01	+ I +
30	.02	+ I +
31	-.09	+XXI +
32	.02	+ I +
33	-.04	+ XI +
34	-.17	X+XXI +
35	.05	+ IX +
36	-.05	+ XI +

Figure 3.5

PROFIL PRÉVISIONNEL AVEC HORIZON DE 12 MOIS



Légende : - : Limite supérieure d'intervalle de confiance prévisionnel de 95 p. 100.
- : Limite inférieure d'intervalle de confiance prévisionnel de 95 p. 100.
+ : Prévision ponctuelle
* : Valeur effective

4. PRÉVOIR AVEC UN MODÈLE À FONCTION DE TRANSFERT

Nous avons vu dans le chapitre précédent que les prévisions sur une série temporelle au moyen d'un modèle ARIMA univarié sont fondées uniquement sur l'historique de la série. Or dans de nombreux cas la série à prévoir pourra être systématiquement influencée par d'autres séries. Dans de telles circonstances on pourra réaliser des prévisions plus précises en tenant compte également de l'historique d'autres séries liées.

Le « modèle à fonction de transfert » est une extension du modèle ARIMA dans laquelle on considère conjointement deux séries ou plus. Dans ce chapitre, nous (1) décrivons brièvement un modèle à fonction de transfert des mises en chantier au Canada que nous avons élaboré et (2) examinons les avantages et désavantages du modèle à fonction de transfert comme outil prévisionnel. On trouvera des analyses des modèles à fonction de transfert dans de nombreux ouvrages, notamment ceux de Box et Jenkins (1970), Granger et Newbold (1977) et Makridakis et al.

4.1 UN MODÈLE À FONCTION DE TRANSFERT DES MISES EN CHANTIER

A. Le modèle du bruit de la fonction de transfert

Soient Y_t les mises en chantier et X_t le prix des logements, tous deux à l'instant t. Pour parvenir à une formulation réaliste de la relation entre les mises en chantier et le prix des logements, il pourra être nécessaire d'inclure dans le modèle de nombreuses valeurs retardées du prix des logements. Car un changement dans ce dernier pourra se répercuter sur les mises en chantier pendant une longue période de temps, d'où la nécessité d'inclure les valeurs retardées. Si une telle influence retardée persiste à se manifester, on considérera le nombre actuel des mises en chantier comme la somme des effets des prix des logements présents et passés. On pourra représenter comme suit une relation dynamique entre les deux variables :

$$Y_t = v_0 X_t + v_1 X_{t-1} + \ldots + v_m X_{t-m} + e_t = v(B) X_t + e_t \qquad (4.1)$$

où $v(B) = vs_o + v_1 B + \ldots + v_m B^m$, les v_j sont des paramètres fixes et les e_t sont des erreurs aléatoires non corrélées dans la série de moyenne nulle et de variance σ^2. Dans la littérature sur les séries temporelles on désigne souvent X_t et T_t séries d'« entrée » et de « sortie », respectivement, et le polynôme $v(B)$ fonction de transfert. Les v_j dans $v(B)$ sont nommés pondérations des réponses impulsionnelles.

En pratique, il se peut que le retard maximum m doive être de valeur assez élevé, de manière à fournir une représentation adéquate de la relation entre les deux variables X et Y. On a proposé plusieurs suggestions à propos des contraintes possibles à imposer à la structure de retard, ce qui permettrait par le fait même de parvenir à une représentation plus économe. On a ainsi suggéré de représenter approximativement le polynôme v(B) par le rapport de deux polynômes en B, et d'écrire (4.1) de la manière suivante :

$$Y_t = \frac{\omega(B)}{\delta(B)} X_t + e_t \qquad (4.2)$$

où

$$\omega(B) = \omega_0 + \omega_1 B + \ldots + \omega_s B^s$$

et

$$\delta(B) = 1 - \delta_1 B - \ldots - \delta_r B^r$$

On parle ici du modèle rationnel à retards échelonnés. Par analogie avec le concept de stationnarité d'un modèle ARMA, on dira que le modèle en (4.2) est « stable » si toutes les racines du polynôme $\delta(B)$ ont une valeur absolue supérieure à un. Cette condition a pour fonction d'assurer que les valeur de X_t dans le lointain passé n'ont qu'une influence négligeable sur Y_t.

Supposons que le terme d'erreur e_t en (4.2) soit un processus ARMA (p, q) de la forme

$$\phi(B) e_t = \theta(B) u_t, \qquad (4.3)$$

où u_t est un processus de type bruit blanc et $\phi(B)$ et $\theta(B)^t$ sont tels que définis dans les modèles ARIMA. En substituant (4.3) dans (4.2), nous pouvons récrire le modèle

$$Y_t = \frac{\omega(B)}{\delta(B)} X_t + \frac{\theta(B)}{\phi(B)} u_t. \qquad (4.4)$$

On parle du <u>modèle à fonction de transfert et bruit blanc</u>, ou plus simplement du <u>modèle à fonction de transfert</u>. Pour de nombreuses applications, les séries en Y_t, X_t et e_t pourront constituer des processus intégrés.

Étant donné que les mises en chantier et les prix des logements sont de caractère saisonnier, avec une période s de douze mois, il faudra, pour élaborer un modèle à fonction de transfert saisonnier, inclure des opérateurs saisonniers dans le bruit. De même que pour les modèles univariés, il s'avère nécessaire de différencier la série de manières saisonnière et non saisonnière pour obtenir la stationnarité de e_t. La fonction de transfert v(B) peut en outre inclure des composantes

saisonnières. On peut donc donner à (4.4) la forme étendue suivante :

$$Y_t = \frac{\omega(B)\Omega(B^s)}{\delta(B)\Delta(B^s)} X_t + \frac{\theta(B)\Theta(B^s)}{\phi(B)\Phi(B^s)} U_t. \qquad (4.5)$$

Nous décrivons maintenant les résultats fournis par le modèle à fonction de transfert des mises en chantier au Canada que nous avons élaboré, la série des prix du logement constituant une variable d'entrée[6]. Cette formulation du modèle à fonction de transfert est compatible avec le paradigme du logement exposé au Chapitre 2; à court terme, la demande de stock agit sur l'offre fixe existante de stock, ce qui détermine le prix, et l'industrie du logement réagit aux prix par de nouvelles constructions.

Nous avons utilisé le nombre mensuel de mises en chantier de maisons individuelles dans les centres urbains canadiens de 10 000 habitants ou plus (indice CANSIM D849796) et les nouveaux indices des prix des logements (indice CANSIM D636200) comme séries d'entrée et de sortie, respectivement. Puisque la série d'indices des prix ne remonte qu'à 1981, les données d'échantillonnage ne consistent qu'en 108 observations mensuelles allant de janvier 1981 à décembre 1989. Tout comme au Chapitre 3 à propos de la modélisation ARIMA, nous avons mis de côté 12 observations portant sur 1989 pour des prévisions post-échantillonnage, les 96 observations restantes servant à la modélisation avec fonction de transfert. Il est possible que la longueur des séries soit insuffisante pour une modélisation efficace, car les deux séries sont marquées par des variations saisonnières.

---

[6] Nous avons également tenté d'élaborer un modèle à fonction de transfert des mises en chantier au Canada en utilisant les traditionnels taux d'intérêt hypothécaires sur 5 ans comme variable d'entrée, les données s'échelonnant de 1981 à 1989. Les résultats indiquent que les taux d'intérêt hypothécaires n'influent pas sur les mises en chantier. Cette constatation concorde avec celle de l'étude Stansell-Mitchell (1985) des marchés du logement américains. Les auteurs ont analysé les relations causales entre six variables différentes décrivant les taux et échéances hypothécaires et les mises en chantier de maisons individuelles aux États-Unis. Leurs résultats indiquent que ni le rationnement du crédit, ni les taux hypothécaires n'agissent sur les mises en chantier. Mais certains indices laissent supposer que les prix des logements sont liés de manière causale aux mises en chantier, conséquence possible de la spéculation. Parmi les autres travaux sur la modélisation avec fonction de transfert du marché du logement figurent ceux de Hillmer et Tiao (1979), Wang et Ma (1981) et Puri et Van Lierop (1988).

de vérification diagnostique en modélisation avec fonction de transfert. L'identification du modèle est plus difficile avec le modèle à fonction de transfert qu'avec le modèle ARIMA, et les outils analytiques nécessaires pour l'identification sont différents dans les deux cas.

L'identification comporte les étapes suivantes : (1) estimation de la fonction de transfert v(B), (2) identification du modèle ARIMA pour déterminer le terme de bruit e_t et (3) détermination des polynômes $\delta(B)$ et $\omega(B)$.

On peut estimer les pondérations des réponses impulsionnelles de la fonction de transfert v(B) de deux façons différentes. Box et Jenkins (1970) ont proposé l'élaboration d'un modèle ARIMA univarié de la forme

$$X_t = \frac{\theta_x(B)}{\phi_x(B)}\alpha_t = T(B)\alpha_t \qquad (4.6)$$

pour la variable d'entrée X_t. On inverse alors le modèle, c'est-à-dire, $T^{-1}(B)X_t = \alpha_t$, et l'on obtient la série résiduelle « préblanchie » α_t. On applique alors le filtre de préblanchiment $T^{-1}(B)$ à la série de sortie Y_t pour obtenir la série filtrée $\beta_t = T^{-1}(B)Y_t$. La série filtrée n'est pas nécessairement du bruit blanc.

La <u>fonction de corrélation croisée</u> (FCC) entre α_t et β_t est proportionnelle à v(B). On peut donc obtenir une estimation de v(B) au moyen de la FCC d'échantillonnage à partir de α_t, des termes résiduels du modèle ARMA ajusté de X_t et β_t, et de la série filtrée Y_t par le T(B) estimé. Bien que la valeur de $\hat{v}(B)$ ainsi obtenue ne constitue pas une estimation précise, elle fournit une point de départ pour l'identification de la fonction de transfert v(B). Nous passons maintenant à la modélisation ARIMA des mises en chantier.

B. Modélisation ARIMA univariée des mises en chantier

L'inspection visuelle du tracé de la série de la série temporelle à la Figure 4.1 révèle que la série des mises en chantier est non stationnaire. Nous avons exprimé la série en logarithmes vulgaires pour obtenir la stationnarité de la variance. Le tracé de la série logarithmique et la FAC de l'échantillon indiquent que des différenciations régulière et saisonnière d'ordre 1 chacune sont nécessaires pour obtenir la stationnarité de la moyenne.

Les FAC et FACP d'échantillon de la série incitent à faire usage d'une MA(1) non saisonnière comme modèle d'essai.

$$(1-B)(1-B^{12})LOGSHS_t = (1-\Theta_{12}B^{12})u_t \qquad (4.7a)$$

où $LOGSHS_t$ est le logarithme vulgaire des mises en chantier de maisons individuelles à l'instant t. Le modèle a été estimé à la fois avec et sans terme d'intersection, le second ayant été choisi parce que ce terme était de valeur insignifiante.

On trouvera au Tableau 4.1(A) les résultats estimés. Le modèle ajusté s'écrit

$$(1-B)(1-B^{12})LOGSHS_t = (1 - 0{,}9548_{12}B^{12})u_t, \quad (4.7b)$$

La valeur du paramètre de mobilité de la moyenne est estimée à 0,9548, ce qui est très proche de la limite du domaine de non-inversibilité[7]. La vérification diagnostique fondée sur les termes résiduels a fourni des résultats satisfaisants, exception faite d'une autocorrélation et d'une autocorrélation partielle significatives de retard 7.

Nous avons procédé à une vérification plus poussée du modèle Ma(1) saisonnier par surajustement et avons obtenu le modèle ARMA(1,1) saisonnier suivant :

$$(1+0.4134B^{12})(1-B)(1-B^{12})LOGSHS_t = (1-0.3765B^{12})u_t, \quad (4.8)$$

où R^2 = 0,936 et SEE = 0,05484. Les deux estimations de coefficient paraissent plausibles et sont statistiquement significatives. Aucune autocorrélation ni autocorrélation partielle n'est significative, bien qu'elles soient considérables aux retards 3 et 7. Quoique le modèle apparût adéquat, ses prévisions ont été moins exactes pour les douze mois de 1989 que celles du modèle MA(1) saisonnier. Les prévisions de ce dernier modèle ont été reproduites au Tableau 4.4(B); elles ont servi de point de comparaison pour vérifier l'efficacité prévisionnelle du modèle à fonction de transfert des mises en chantier.

---

[7] Pour l'estimation d'un modèle ARIMA les programmes TSP et SHAZAM utilisent une méthode ML complète en procédant à ce qu'il est convenu d'appeler la rétroconversion, alors que le programme SCA-UTS fait usage d'une méthode ML approximative, malgré qu'on puisse utiliser en option sa méthode « exacte » à des fins d'estimation. Les différences quant à l'estimation des mêmes paramètres par des programmes d'ordinateur différents pour le même modèle (4.7) et les mêmes données sont remarquables : $\hat{\Theta}_{12}$ et son erreur-type égalent respectivement 0,9548 et 0,0739 pour SCA-UTS, 0,8595 et 0,0390 pour TSP et 0,8598 et 0,0376 pour SHAZAM.

C. Modélisation ARIMA univariée des indices des prix des logements

L'inspection visuelle du tracé de la série temporelle de la Figure 4.2 révèle que la série des indices des prix des logements est non stationnaire. Pour obtenir la stationnarité de la série nous avons calculé son logarithme vulgaire et avons procédé deux fois à une différenciation régulière de la série logarithmique.

À partir de l'examen des FAC et FACP d'échantillon des séries différenciées, nous avons conclu que la série d'entrée constitue un modèle ARIMA (1, 2, 1) régulier :

$$(1-\phi_1 B)(1-B)^2 PRICE_t = (1-\theta_1 B)\alpha_t \qquad (4.9a)$$

où $PRICE_t$ est le logarithme vulgaire de l'indice d'origine du prix des logements.

On trouvera résumées au Tableau 4.2 les estimations ML du modèle en (4.9a). Le modèle ajusté est

$$(1-0.5899B)(1-B)^2 PRICE_t = (1-0.9010B)\alpha_t. \qquad (4.9b)$$

Les estimations des deux paramètres sont significatives et l'ajustement est très bon, comme l'indique la valeur $R^2 = 0,998$. Les autocorrélations et autocorrélations partielles sont significatives aux retards 6 et 16.

Le modèle ARIMA multiplicatif suivant est un des nombreux modèles additionnels qui ont été estimés :

$$(1-0.6768B)(1+0.6038B^{12})(1-B)(1-B^{12})PRICE_t = \alpha_t. \qquad (4.10)$$

Les deux estimations de paramètre sont significatives et satisfont aux conditions de stationnarité. L'ajustement est très bon si l'on en juge par la valeur $R^2 = 0,998$. Les autocorrélations sont importantes, mais non significatives, à de nombreux retards et l'autocorrélation partielle est non significative au retard 12. Mais son efficacité prévisionnelle postéchantillonnage pour les 12 mois de 1989 est inférieure à celle du modèle en (4.9). Aussi utilise-t-on le modèle ARIMA (1, 2, 1) pour filtrer les séries d'entrée et de sortie.

D. Modélisation avec fonction de transfert des mises en chantier

La FCC de l'échantillon de la série d'entrée pré-blanchie et la série de sortie filtrée n'ont pas la clarté nécessaire pour l'identification de la fonction de transfert. Nous avons donc eu recours à une méthode d'identification de rechange.

On peut estimer le modèle linéaire (4.1) avec une valeur relativement élevée de m au moyen de la méthode des moindres carrés. Étant donné que e_t n'est pas nécessairement un processus de bruit blanc, Liu et Hanssens (1982) proposent l'inclusion d'un modèle à termes de perturbation autorégressifs de la forme $(1-\phi_1 B)e_t = u_t$ ou $(1-\phi_1 B)(1-\Phi_1 B^s)e_t = u_t$ pour améliorer la précision de l'estimation des moindres carrés des pondérations des réponses impulsionnelles. Ils nomment cette approche méthode à fonction de transfert linéaire. En utilisant la méthode de Liu et Hanssens, nous avons à titre d'essai identifié comme suit un modèle à fonction de transfert :

$$(1-B)(1-B^{12})\,LOGSHS_t = (v_0 + v_1 B + \ldots + v_4 B^4)(1-B)^2\,PRICE_t$$

où les estimations préliminaires des v sont $\hat{v} = 2{,}38$, $\hat{v}_1 = 0{,}49$, $\hat{v}_2 = 0{,}69$, $\hat{v}_3 = -0{,}63$ et $\hat{v}_4 = 2{,}83$°. Seules deux estimations, celles de $_o$ et $_4$, sont significatives.

Disposant d'une estimation de v(B), on peut identifier un modèle ARMA de la composate bruit en examinant les FAC et FACP d'échantillon de la série estimée du bruit :

$$\hat{e}_t = Y_t - \hat{v}(B)X_t,$$

où $Y_t = (1-B)(1-B^{12})\,LOGSHS_t$ et $X_t = (1-B)^2\,PRICE_t$. Nous avons en guise d'essai identifié le modèle ARIMA multiplicatif

$$e_t = (1-\theta_1 B)(1-\Theta_1 B^{12})\,u_t$$

au processus de bruit.

Une fois les ordres de la fonction de transfert et de la composante bruit identifiés, on peut estimer efficacement le modèle pleinement identifié en maximisant la fonction de vraisemblance correspondante. Lorsque le modèle identifié a été estimé au moyen de la méthode ML, les statistiques de t pour les trois paramètres du numérateur $\hat{\omega}_1$, $\hat{\omega}_2$ et $\hat{\omega}_3$ n'étaient pas significatives. Ces paramètres peuvent être supprimés du modèle en occasionnant une perte minime de pouvoir explicatif.

Le modèle à fonction de transfert à estimer est alors

$$(1-B)(1-B^{12})\,LOGSHS_t = (\omega_0 + \omega_4 B^4)(1-B)^2\,PRICE_t + (1-\theta_1 B)(1-\Theta_1 B^{12})\,u_t \quad (4.11a)$$

On trouvera au Tableau 4.2 les résultats de l'estimation de ce modèle. Le modèle ajusté est

$$(1-B)(1-B^{12})\,LOGSHS_t = (2{,}3845 + 2{,}8852 B^4)(1-B)^2\,PRICE_t + (1+0{,}2138B)(1-0{,}94431 B^{12})\,u_t. \quad (4.11b)$$

Notons que toutes les estimations de paramètre sont significatives et que l'ajustement du modèle est bon si l'on en juge par la valeur de R^2, soit 0,938.

On a vérifié l'adéquation du modèle estimé. Les termes résiduels d'un modèle adéquat devraient se comporter comme une série de bruit blanc. Les FAC et FACP d'échantillon des termes résiduels et les tests de Ljung-Box Q sont des instruments utiles pour la vérification diagnostique. La FCC d'échantillon entre les termes résiduels et la série en X_t préblanchie ($\hat{a}_t$) est également un outil diagnostique utile. Si le modèle est adéquat, les corrélations croisées d'échantillon devraient être significatives pour tous les retards.

La FCC d'échantillon entre les termes résiduels et l'entrée préblanchie est donnée au Tableau 4.3 et tracée à la Figure 4.3. Aucune corrélation croisée n'est significative au niveau de signification 0,05, ce qui indique que l'indice du prix des logements est une sortie exogène comme requis. On trouvera aux Figures 4.4(A) et (B) respectivement les FAC et FACP d'échantillon des termes résiduels. Aucun coefficient n'est significatif, et les statistiques Ljung-Box Q concordent avec l'hypothèse que les termes résiduels sont du bruit blanc. L'analyse des FAC, FACP et FCC d'échantillon n'a pas fait ressortir de structure restante marquée.

On peut approfondir la compréhension du modèle en examinant la matrice de corrélation des estimations des paramètres. Bien que des corrélations élevées puissent être la conséquence de la série précise de données modélisées, elles peuvent également être compatibles avec un défaut de spécification du modèle ou avec une redondance des paramètres. Nous avons observé que tous les coefficients de corrélation sont faibles, le plus fort, entre $_0$ et $_4$, étant égal à 0,25. Nous concluons que le modèle à fonction de transfert ajusté est adéquat.

Pour conclure, nous allons examiner pourquoi il est avantageux d'utiliser pour les mises en chantier un modèle à fonction de transfert au lieu d'un modèle univarié. Le modèle univarié résumé au Tableau 4.1(A) est

$$(1-B)(1-B^{12})LOGSHS_t = (1-0.9548B^{12})u_t$$

Cela est comparable en structure à la composante bruit du modèle à fonction de transfert

$$c_t = (1+0.2138B)(1-0.9443B^{12})u_t.$$

L'erreur-type résiduelle pour le modèle ARIMA est de ,05678 alors que celle du modèle à fonction de transfert est de ,05371. Malgré que la réduction de l'erreur-type résiduelle soit minime (quelque 5 p. 100), le modèle à fonction de transfert n'en

explique pas moins une proportion plus élevée de la variabilité des données sur les mises en chantier.

4.2 PRÉVOIR AVEC UN MODÈLE À FONCTION DE TRANSFERT

Une fois obtenu un modèle à fonction de transfert adéquat, on peut l'utiliser pour prévoir les valeurs futures de la variable de sortie. Même si nous avons décrit brièvement plus bas la procédure à suivre pour faire des prévisions au moyen d'un modèle à fonction de transfert avec des variables d'entrée et de sortie stationnaires, la même procédure devrait s'appliquer au modèle à fonction de transfert saisonnier multiplicatif avec des variables d'entrée et de sortie intégrées.

Supposons pour l'instant que le modèle soit parfaitement connu et que nous ayons besoin à l'instant T d'une prévision de Y_t à h étapes vers l'avenir. Le modèle à fonction de transfert « connu » est de la forme

$$Y_t = \frac{\omega(B)}{\delta(B)} X_t + \frac{\theta(B)}{\phi(B)} u_t \qquad (4.12a)$$

et peut être écrit

$$\delta^*(B) Y_t = \omega^*(B) X_t + \theta^*(B) u_t, \qquad (4.12b)$$

où $\delta^*(B)$, $\omega^*(B)$ et $\theta^*(B)$ sont des polynômes en B. Si le modèle est valide pour l'avenir, on peut, à partir de (4.12b), écrire Y_{T+h} comme suit :

$$\begin{aligned} Y_{T+h} = {} & \delta_1^* Y_{T+h-1} + \cdots + \delta_{r^*}^* Y_{T+h-r^*} + \omega_0^* X_{T+h} - \cdots - \omega_{s^*}^* X_{T+h-s^*} \\ & + u_{T+h} - \theta_1^* u_{T+h-1} - \cdots - \theta_{g^*}^* u_{T+h-g^*}. \end{aligned} \qquad (4.13)$$

Comme dans le cas du modèle ARIMA, l'espérance conditionnelle de Y_{T+h} étant donné l'ensemble d'information I_t = (Y_T, Y_{T-1}, ...; X_T, X_{T-1}, ...) correspond à la prévision à erreur quadratique minimale à l'instant T si les séries u_t sont normales. On peut donc calculer la prévision optimale à h étapes vers l'avenir en reprenant l'espérance conditionnelle de (4.13) :

$$\begin{aligned} f_T(h) = E_T(Y_{T+h}) = {} & \delta_1^* E_T(Y_{T+h-1}) + \cdots + \delta_{r^*}^* E_T(Y_{T+h-r^*}) + \omega_0^* E_T(X_{T+h}) \\ & - \cdots - \omega_{s^*}^* E_T(X_{T+h-s^*}) + E_T(u_{T+h}) - \theta_1^* E_T(u_{T+h-1}) - \cdots - \theta_{g^*}^* E_T(u_{T+h-g^*}), \end{aligned} \qquad (4.14)$$

où $f_T(h)$ représente la prévision à h étapes vers l'avenir et $E_T[.]$ l'espérance conditionnelle, toutes deux à l'instant T.

Notons que (4.14) renferme l'espérance conditionnelle de la série X_t. La prédiction de Y_t requiert donc celle de X_t.

Puisque la série X_t est un processus ARIMA, on peut calculer ses espérances ou prévisions conditionnelles, $E_T(X_{T+h}) = X_{T+h}$, tel que décrit par (3.9) à la section 3.2.

Si les processus X_t et u_t sont normaux, l'erreur $e_T(h)$ pour une prévision à h étapes vers l'avenir est normale, et l'on peut

obtenir une prévision à intervalle. Ainsi, une prévision à intervalle de 95 p. 100 sera donnée par

$$f_T(h) \pm 1.96\, SE[e_T(h)], \tag{4.15}$$

où $SE(e_T(h))$ est l'erreur-type de prévision et dépend des paramètres du modèle de même que des variances de X_t et u_t. À mesure que s'éloigne l'horizon prévisionnel h, la variance prévisionnelle en (4.19) s'accroît de même que la largeur d'une prévision à intervalle.

Le modèle à fonction de transfert ajusté est

$$\begin{aligned}(1-B)(1-B^{12})LOGSHS_t = {} & (2.3845 + 2.8852B^4)(1-B)^2 PRICE_t \\ & + (1 + .2138B)(1 - .9443B^{12})u_t\end{aligned}$$

et le modèle d'entrée est

$$(1 - .5899B)(1-B)^2 PRICE_t = (1 - .9010B)\alpha_t.$$

En résolvant par multiplication le modèle à fonction de transfert, nous obtenons la fonction de prévision à h étapes vers l'avenir à l'origine T suivante :

$$\begin{aligned}E_T(LOGSHS_{T+h}) = {} & E_T(LOGSHS_{T+h-1}) + E_T(LOGSHS_{T+h-12}) \\ & - E_T(LOGSHS_{T+h-13}) + 2.3845E_T(PRICE_{T+h}) \\ & - 4.7690E_T(PRICE_{T+h-1}) + 2.3845E_T(PRICE_{T+h-2}) \\ & + 2.8852E_T(PRICE_{T+h-4}) + E(u_{T+h}) \\ & + .2138E(u_{T+h-1}) - .9443E(u_{T+h-12}) \\ & + .2019E(u_{T+h-13}).\end{aligned} \tag{4.16}$$

où $E_T[.]$ représente l'espérance conditionnelle. Les espérances conditionnelles de la série des prix sont obtenues à partir de son modèle ARIMA univarié sous la forme

$$\begin{aligned}E_T(PRICE_{T+h}) = {} & 2.5899E_T(PRICE_{T+h-1}) + 2.1798E_T(PRICE_{T+h-2}) \\ & - .5899E_T(PRICE_{T+h-3}) - E_T(\alpha_{T+h}) \\ & - .9018E_T(\alpha_{T+h-1}).\end{aligned} \tag{4.17}$$

On trouvera au Tableau 4.4(A) les prévisions et leurs statistiques associées tirées de (4.16) pour des horizons h = 1, 2, ..., 12. Il s'agit des prévisions faites en décembre 1988 pour les 12 mois de 1989. On a inscrit dans les trois colonnes suivantes les erreurs-types des prévisions, les valeurs effectives et les erreurs de prévision. Les erreurs de prévision exprimées en pourcentage des valeurs effectives varient en valeur absolue de 0,446 pour octobre 1989 à 4,503 pour juin 1989. Si les points d'inflexion ont été relativement bien prédits, les prévisions n'en dépassent pas moins systématiquement les valeurs effectives. Il convient de noter que pour le calcul de ces prévisions des mises en chantier, nous avons utilisé comme

données d'entrée les prévisions de l'indice du prix des logements tirées du modèle ARIMA (1,2,1) en (4.17)

À des fins de comparaison, nous avons effectué des prévisions de mises en chantier s'étalant sur douze mois au moyen du modèle ARIMA univarié en (4.7) :

$$(1-B)(1-B^{12})LOGSHS_t = (1-.9548B)u_t.$$

Nous avons reproduit au Tableau 4.4(B) ces prévisions et leurs statistiques associées.

Définissons l'erreur absolue moyenne en pourcentage (MAPE) et la moyenne quadratique des erreurs en pourcentage (RMSPE) respectivement comme

$$MAPE = \frac{\sum_{i=1}^{n} 100|(A_t - F_t)/A_t|}{n}$$

et

$$RMSPE = \sqrt{\frac{\sum_{i=1}^{n} [100(A_t - F_t)/A_t]^2}{n}},$$

où F_t et A_t sont les valeurs prédites et effectives et n le nombre de prévisions. Les MAPE sont respectivement de 19,22 et 12,50 pour les prévisions avec fonction de transfert et avec modèle ARIMA, les RMSPE valant respectivement 56,89 et 42,95. Manifestement, le modèle à fonction de transfert n'a pas fourni d'aussi bonnes prévisions que le modèle ARIMA.

On note au Tableau 4.4(B) que les erreurs-types prévisionnelles sont plus grandes avec le modèle à fonction de transfert qu'avec le modèle univarié pour tous les temps d'horizon sauf $h = 1$. On s'attendrait à ce que le modèle à fonction de transfert fournisse des prévisions à court terme plus précises s'il était tenu compte de l'information supplémentaire contenue dans la série d'entrée. Il faut toutefois faire usage des prévisions fondées sur les données d'entrée plutôt que des observations effectives lorsqu'on a affaire à des temps d'horizon plus longs.

Les prévisions fournies par un modèle à fonction de transfert devraient être plus précises que celles fournies par un modèle univarié (1) si les variables d'entrée expliquent une proportion importante des variations de la variable de sortie et (2) si les variables d'entrée peuvent être prédites avec précision. Quoique la fonction de transfert ajustée (4.16) explique une proportion plus élevée des mises en chantier que le modèle ARIMA univarié, l'influence du prix des logements sur les mises en chantier ne paraît pas assez forte pour permettre une amélioration des prévisions par la modélisation à fonction de transfert.

En tant qu'outil prévisionnel, le modèle à fonction de transfert présente les avantages suivants :

(1) il fournit des prévisions précises tout au moins à court terme;
(2) en utilisant des informations contenues dans d'autres séries liées, le modèle à fonction de transfert peut procurer des prévisions plus précises que les modèles ARIMA univariés; et
(3) et il est possible d'estimer un modèle à fonction de transfert en utilisant des données remontant à tout instant dans le passé et d'obtenir avec un tel modèle des prévisions pour tout horizon temporel, ce qui permet de vérifier l'efficacité prévisionnelle du modèle.

Il faut en revanche noter quelques inconvénients :

(1) la prévision avec une fonction de transfert nécessite la prédiction des variables d'entrée; or les erreurs de prévision des variables d'entrée peuvent rendre les prévisions du modèle à fonction de transfert très imprécises; et
(2) tout comme le modèle économétrique traditionnel, le modèle à fonction de transfert exige que les variables d'entrée soient exogènes. Si les variables d'entrée ne sont pas exogènes, la modélisation et la prévision avec fonction de transfert ne sont pas valides.

Tableau 4.1

(A) RÉSUMÉ D'UN MODÈLE ARIMA DES MISES EN CHANTIER
(Maisons individuelles en logarithmes vulgaires, d = D = 1)

VARIABLE	TYPE DE VARIABLE	ORIGINEL OU CENTRE	DIFFÉRENCIATION
LNSHS	ALÉATOIRE	ORIGINEL	$(1-B^{1})(1-B^{12})$

DÉSIGNATION DU PARAMÈTRE	NOM DE LA VARIABLE	VALEUR	ERREUR DÉVIATION STANDARD	VALEUR T
1 THETA1	LOGHS	0.9548	0.0739	12.92

SOMME TOTALE DES CARRÉS	0.449533E+01
NOMBRE TOTAL D'OBSERVATIONS	96
SOMME RÉSIDUELLE DES CARRÉS	0.257596E+00
R AU CARRÉ	0.931
NOMBRE EFFECTIF D'OBSERVATIONS	83
ESTIMATION DE LA VARIANCE RÉSIDUELLE	0.322404E-02
ERREUR-TYPE RÉSIDUELLE	0.567807E-01

(B) RÉSUMÉ D'UN MODÈLE ARIMA DES INDICES DU PRIX DES LOGEMENTS
(Base = 1981 en logarithmes vulgaires, d = 2)

VARIABLE	TYPE DE VARIABLE	ORIGINEL OU CENTRE	DIFFÉRENCIATION
LOGHS	ALÉATOIRE	ORIGINEL	$(1-B^{1})(1-B^{1})$

DÉSIGNATION DU PARAMÈTRE	NOM DE LA VARIABLE	VALEUR	ERREUR DÉVIATION STANDARD	VALEUR T
1 THETA1				
2 PHI1				

SOMME TOTALE DES CARRÉS	0.126748E+01
NOMBRE TOTAL D'OBSERVATIONS	96
SOMME RÉSIDUELLE DES CARRÉS	0.188733E-02
R AU CARRÉ	0.998
NOMBRE EFFECTIF D'OBSERVATIONS	93
ESTIMATION DE LA VARIANCE RÉSIDUELLE	0.202938E-04
ERREUR-TYPE STANDARD	0.450487E-02

Tableau 4.2

(A) RÉSUMÉ D'UN FONCTION DE TRANSFERT DES MISES EN CHANTIER
(Maisons individuelles en logarithmes vulgaires, d = D = 1)

VARIABLE	TYPE DE VARIABLE	ORIGINEL OU CENTRE	DIFFÉRENCIATION
LMSHS	ALÉATOIRE	ORIGINEL	$(1-B^{1})$ $(1-B^{12})$
PRICE	ALÉATOIRE	ORIGINEL	$(1-B^{1})$ $(1-B^{1})$

DÉSIGNATION DU PARAMÈTRE		NOM DE LA VARIABLE	VALEUR	ERREUR DÉVIATION STANDARD	VALEUR T
1	W0	PRICE	2.3845	0.9157	2.60
2	W4	PRICE	2.8852	1.0610	2.72
3	THETA1	LOGSHS	-0.2138	0.1079	-1.98
4	THETA12	LOGSHS	0.9443	0.0732	12.90

SOMME TOTALE DES CARRÉS	0.449533 E+01
NOMBRE TOTAL D'OBSERVATIONS	96
SOMME RÉSIDUELLE DES CARRÉS	0.239441 E+00
R AU CARRÉ	0.938
NOMBRE EFFECTIF D'OBSERVATIONS	83
ESTIMATION DE LA VARIANCE RÉSIDUELLE	0.288484 E-02
ERREUR-TYPE RÉSIDUELLE	0.537107 E-01

Tableau 4.3

CORRÉLATIONS CROISÉES ENTRE LES INDICES DES PRIX
PRÉBLANCHIS ET LES RÉSIDUS DE LA FONCTION DE TRANSFERT

PÉRIODE DE TEMPS ANALYSÉE	14 to 96	
NOM DE LA SÉRIE	RX	RY
NOMBRE EFFECTIF D'OBSERVATIONS	83	83
DÉVIATION STANDARD DE LA SÉRIE	0.0044	0.0486
MOYENNE DE LA SÉRIE (DIFFÉRENCIÉE)	0.0004	0.0040
DÉVIATION STANDARD DE LA MOYENNE	0.0005	0.0053
VALEUR T DE LA MOYENNE (PAR RAPPORT À ZÉRO)	0.9134	0.7536

CORRÉLATION ENTRE RY ET ... RX = 0.02

CORRÉLATION CROISÉE ENTRE RX(T) ET ... RY(T-L)

1-12	-.06	.06	-.13	-.03	-.09	.01	-.13	.10	-.18	-.04	-.10	.04
Er.-T.	.11	.11	.11	.11	.11	.11	.11	.12	.12	.12	.12	.12
13-24	.11	.00	-.09	.11	-.11	.12	.03	-.06	.01	.01	-.04	-.02
Er.-T.	.12	.12	.12	.12	.12	.12	.13	.13	.13	.13	.13	.13
25-36	.02	-.01	.02	-.09	.06	-.02	.02	.03	.01	-.03	-.00	-.04
Er.-T.	.13	.13	.13	.13	.14	.14	.14	.14	.14	.14	.14	.15

CORRÉLATION CROISÉE ENTRE RY(T). ET ... RX(T-L)

1-12	-.07	.15	-.07	-.16	.09	.22	-.08	.07	-.08	-.07	.03	-.00
Er.-T.	.11	.11	.11	.11	.11	.11	.11	.12	.12	.12	.12	.12
13-24	.17	-.00	-.13	.05	.02	-.07	.09	.05	-.08	-.03	-.15	.02
Er.-T.	.12	.12	.12	.12	.12	.12	.13	.13	.13	.13	.13	.13
25-36	-.14	.18	-.00	-.08	-.07	-.10	.05	.10	.09	.05	-.14	.01
Er.-T.	.13	.13	.13	.13	.14	.14	.14	.14	.14	.14	.14	.15

Tableau 4.4

PRÉVISIONS 12 MOIS VERS L'AVENIR DES MISES EN CHANTIER
(En logarithmes vulgaires)

(A) Fondées sur un modèle à fonction de transfert

MOIS	PRÉVISION	ERREUR DÉVIATION STANDARD	VALEUR EFFECTIVE	ERREUR	ERREUR EN POURCENTAGE	ERREUR EN % AU CARRÉ
Jan.	3.824	.0548	3.770	-.054	-1.444	2.086
Fév.	3.737	.0855	3.699	-.038	-1.038	1.078
Mar.	3.842	.1076	3.803	-.040	-1.045	1.093
Avr.	4.038	.1259	4.285	-.064	-1.612	2.600
Mai	4.149	.1428	4.067	-.082	-2.022	4.087
Juin	4.197	.1574	4.016	-.181	-4.503	20.275
Juil.	4.081	.1706	3.985	-.096	-2.406	5.788
Août	4.032	.1828	3.998	-.033	-.836	.699
Sep.	4.011	.1942	3.943	-.068	-1.721	2.961
Oct.	3.952	.2050	3.934	-.018	-.446	.199
Nov.	3.942	.2152	3.902	-.040	-1.030	1.061
Déc.	3.869	.2249	3.827	-.043	-1.118	1.251

(B) Fondées sur un modèle ARIMA

MOIS	PRÉVISION	ERREUR DÉVIATION STANDARD	VALEUR EFFECTIVE	ERREUR	ERREUR EN POURCENTAGE	ERREUR EN % AU CARRÉ
JAN.	3.766	.0568	3.770	.004	.104	.011
FÉV.	3.683	.0803	3.699	.015	.412	.169
MAR.	3.768	.0983	3.803	.035	.922	.849
AVR.	4.013	.1136	3.974	-.039	-.971	.943
MAI	4.141	.1270	4.067	-.073	-1.804	3.254
JUIN	4.119	.1391	4.016	-.103	-2.558	6.542
JUIL.	4.026	.1502	3.985	-.041	-1.020	1.041
AOÛT	4.002	.1606	3.998	-.003	-.081	.007
SEP.	3.974	.1703	3.943	-.031	-.795	.632
OCT.	3.973	.1796	3.934	-.038	-.973	.946
NOV.	3.955	.1883	3.902	-.053	-1.358	1.844
DÉC.	3.884	.1967	3.827	-.057	-1.501	2.252

Figure 4.1 Mises en chantier de maisons individuelles
en logarithmes vulgaires

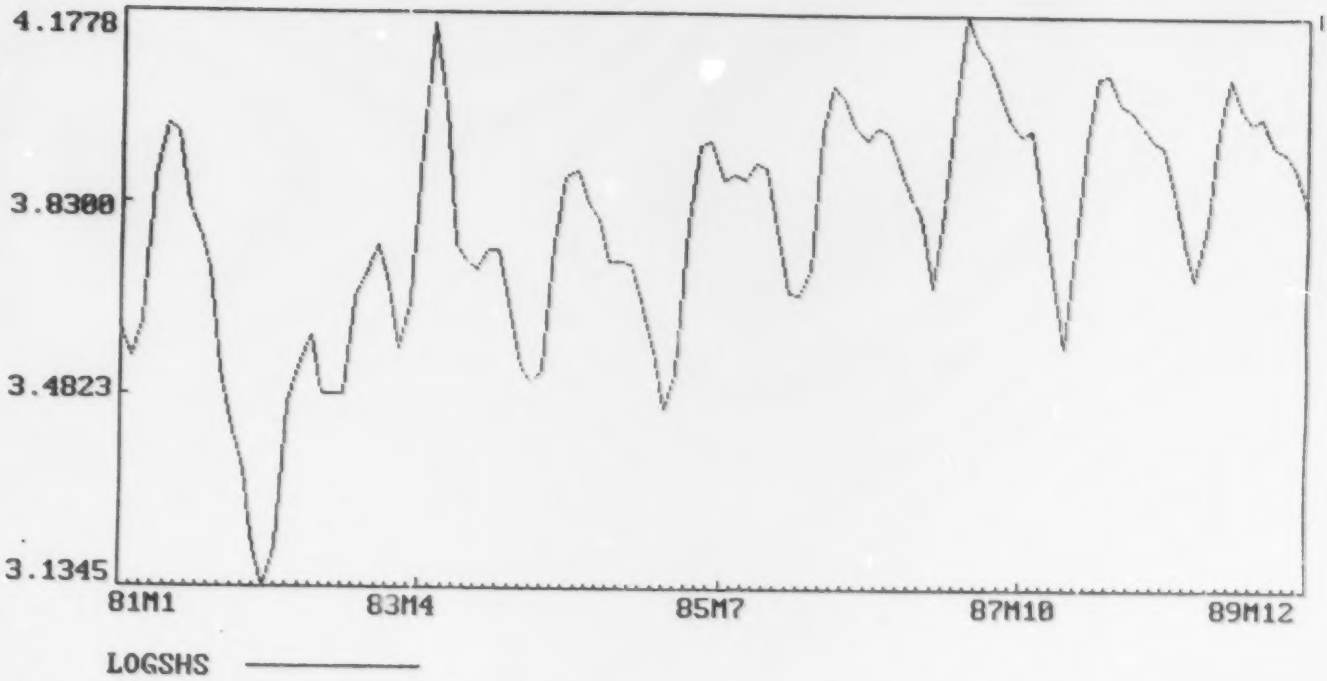


Figure 4.2 Indice du prix des logements
(Base 1981) en logarithmes vulgaires

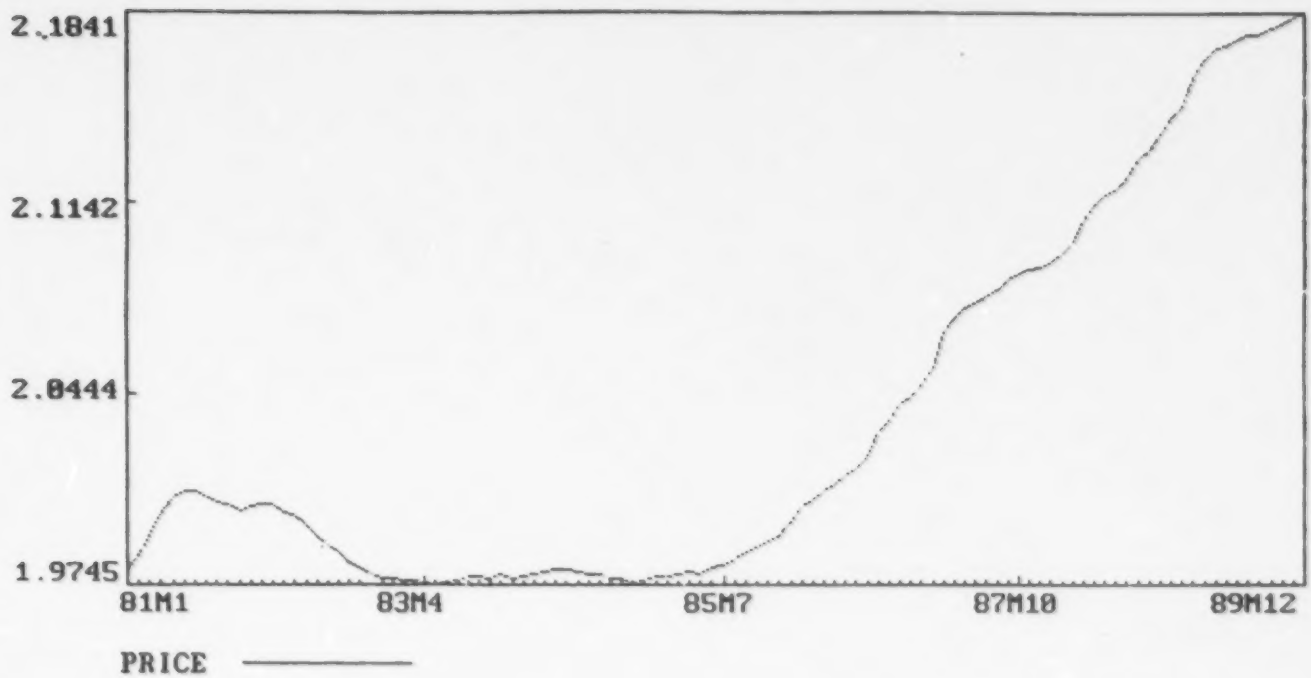


Figure 4.3

CORRÉLATIONS CROISÉES ENTRE LES INDICES DES PRIX PRÉBLANCHIS ET LES RÉSIDUS DE LA FONCTION DE TRANSFERT

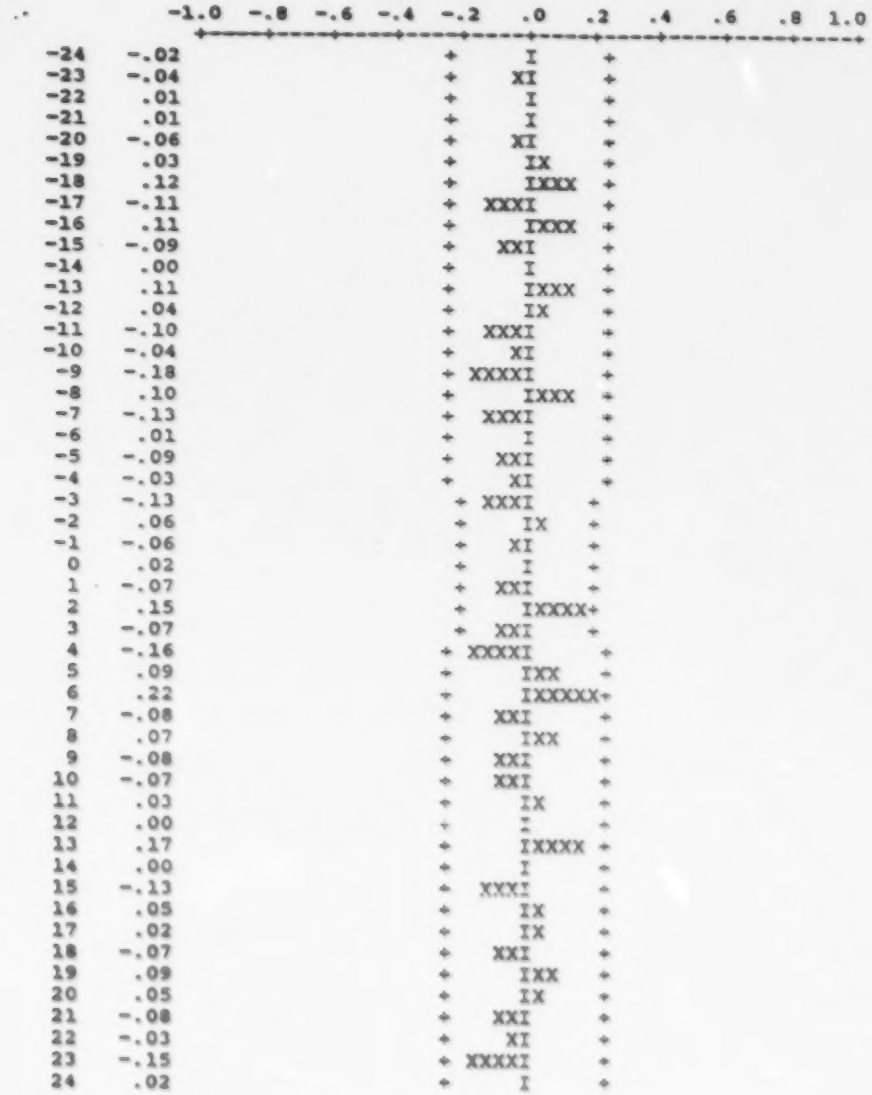


Figure 4.4 (A)

AUTOCORRÉLATIONS D'ÉCHANTILLON DES RÉSIDUS DE LA FONCTION DE TRANSFERT

PÉRIODE DE TEMPS ANALYSÉE	14 to 96
NOM DE LA SÉRIE	RT
NOMBRE EFFECTIF D'OBSERVATIONS	83
DÉVIATION STANDARD DE LA SÉRIE	0.0486
MOYENNE DE LA SÉRIE (DIFFÉRENCIÉE)	0.0040
DÉVIATION STANDARD DE LA MOYENNE	0.0053
VALEUR T DE LA MOYENNE (PAR RAPPORT À ZÉRO)	0.7536

1-12	-.06	-.03	-.08	.12	.11	-.04	-.10	-.10	-.11	.04	.06	-.17
Er.-T.	.11	.11	.11	.11	.11	.11	.11	.12	.12	.12	.12	.12
Q	.4	.4	1.0	2.3	3.4	3.5	4.4	5.3	6.5	6.7	7.1	10.0
13-24	.02	.08	.10	-.07	-.15	-.01	.03	.18	-.00	-.09	-.12	.04
Er.-T.	.12	.12	.12	.12	.12	.13	.13	.13	.13	.13	.13	.13
Q	10.1	10.8	11.9	12.5	15.0	15.0	15.1	18.6	18.6	19.6	21.3	21.5

		-1.0 -.8 -.6 -.4 -.2 .0 .2 .4 .6 .8 1.0
1	-.06	XXI
2	-.03	XI
3	-.08	XXI
4	.12	IXXX
5	.11	IXXX
6	-.04	XI
7	-.10	XXI
8	-.10	XXI
9	-.11	XXXI
10	.04	IX
11	.06	IXX
12	-.17	XXXXI
13	.02	IX
14	.08	IXX
15	.10	IXXX
16	-.07	XXI
17	-.15	XXXXI
18	-.01	I
19	.03	IX
20	.18	IXXXX
21	.00	I
22	-.09	XXI
23	-.12	XXXI
24	.04	IX

Figure 4.4 (B)

AUTOCORRÉLATIONS PARTIELLES D'ÉCHANTILLON DES RÉSIDUS DE LA FONCTION DE TRANSFERT

1-12	-0.6	-.03	-.09	.11	.12	-.03	-.08	-.11	-.17	-.00	.09	-.14
Er.-T.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11
13-24	.06	.12	.05	-.07	-.17	-.09	-.04	.20	.09	.01	-.08	-.11
Er.-T.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11
25-36	-.06	-.12	.00	.04	.06	.13	-.07	.00	.02	-.05	.10	-.10
Er.-T.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11

```
                -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
                 +----+----+----+----+----+----+----+----+----+----+----+
 1    -.06                          +  XXI    +
 2    -.03                          +   XI    +
 3    -.09                          +  XXI    +
 4     .11                          +    IXXX +
 5     .12                          +    IXXX +
 6    -.03                          +   XI    +
 7    -.08                          +  XXI    +
 8    -.11                          + XXXI    +
 9    -.17                          +XXXXI    +
10     .00                          +    I    +
11     .09                          +    IXX  +
12    -.14                          +XXXXI    +
13     .06                          +    IXX  +
14     .12                          +    IXXX +
15     .05                          +    IX   +
16    -.07                          +  XXI    +
17    -.17                          +XXXXI    +
18    -.09                          +  XXI    +
19    -.04                          +   XI    +
20     .20                          +    IXXXXX
21     .09                          +    IXX  +
22     .01                          +    I    +
23    -.08                          +  XXI    +
24    -.11                          + XXXI    +
```

Figure 4.4 (C)

AUTOCORRÉLATIONS D'ÉCHANTILLON DU BRUIT DE LA FONCTION DE TRANSFERT

PÉRIODE DE TEMPS ANALYSÉE	14 to 96
NOM DE LA SÉRIE	NT
NOMBRE EFFECTIF D'OBSERVATIONS	83
DÉVIATION STANDARD DE LA SÉRIE	0.0805
MOYENNE DE LA SÉRIE (DIFFÉRENCIÉE)	0.0043
DÉVIATION STANDARD DE LA MOYENNE	0.0088
VALEUR T DE LA MOYENNE (PAR RAPPORT À ZÉRO)	0.4857

1-12	.24	.02	.04	.18	.22	.02	-.15	-.30	-.13	.05	-.05	-.51
Er.-T.	.11	.12	.12	.12	.12	.12	.12	.13	.13	.14	.14	.14
Q	5.1	5.1	5.3	8.3	12.5	12.6	14.6	22.9	24.5	24.8	25.1	50.6
13-24	-.12	.04	.07	-.05	-.19	-.08	.10	.25	.07	-.13	-.13	.13
Er.-T.	.16	.16	.16	.16	.16	.16	.16	.16	.17	.17	.17	.17
Q	52.1	52.3	52.7	53.0	56.9	57.5	58.6	65.7	66.2	68.1	70.1	72.1
25-36	.04	-.07	-.08	-.07	.10	.11	-.03	-.06	.05	.15	.16	-.10
Er.-T.	.17	.17	.17	.17	.17	.17	.17	.17	.17	.18	.18	.18
Q	72.2	72.8	73.6	74.2	75.7	77.3	77.4	78.0	78.3	81.4	85.1	86.5

```
                -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
                 +----+----+----+----+----+----+----+----+----+----+----+
 1    .24                          +    IXXXX+X
 2    .02                         +     I    +
 3    .04                         +     IX   +
 4    .18                         +     IXXXXX+
 5    .22                         +     IXXXXX+
 6    .02                         +     IX   +
 7   -.15                         + XXXXI    +
 8   -.30                        X+XXXXXI    +
 9   -.13                        -   XXXI     -
10    .05                        +     IX     -
11   -.05                        +    XI      +
12   -.51                 XXXXXX+XXXXXXI      -
13   -.12                       +    XXXI     +
14    .04                       +      IX     +
15    .07                       -      IXX    +
16   -.05                       +     XI      +
17   -.19                       +  XXXXXI      -
18   -.08                       +    XXI      +
19    .10                       +      IXX    +
20    .25                       +      IXXXXXX+
21    .07                       +      IXX    +
22   -.13                       +    XXXI     +
23   -.13                       +    XXXI     +
24    .13                       +      IXXX   +
```

Figure 4.4 (D)

AUTOCORRÉLATIONS D'ÉCHANTILLON DU BRUIT DE LA FONCTION DE TRANSFERT

1-12	.24	-.05	.05	.17	.14	-.06	-.16	-.31	-.10	.09	.01	-.43
Er.-T.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11
13-24	.24	.07	-.00	-.06	-.12	-.05	.09	.01	-.03	.02	-.13	-.18
Er.-T.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11
25-36	-.04	-.06	.13	.04	.05	.05	-.05	.03	.02	-.07	.01	-.13
Er.-T.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11

```
          -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
           +----+----+----+----+----+----+----+----+----+----+----+
 1   .24                      +    IXXXX+X
 2  -.05                      +   XI    +
 3   .05                      +    IX   +
 4   .17                      +    IXXXX+
 5   .14                      +    IXXXX+
 6  -.06                      +   XI    +
 7  -.16                      +XXXXI    +
 8  -.31                   XXX+XXXXI    +
 9  -.10                      +  XXI    +
10   .09                      +    IXX  +
11   .01                      +    I    +
12  -.43                   XXXXXX+XXXXI    +
13   .24                      +    IXXXX+X
14   .07                      +    IXX  +
15   .09                      +    I    +
16  -.06                      +   XI    +
17  -.12                      + XXXI    +
18  -.05                      +   XI    +
19   .09                      +    IXX  +
20   .01                      +    I    +
21  -.03                      +   XI    +
22   .02                      +    I    +
23  -.13                      + XXXI    +
24  -.18                      XXXXXI    +
```

5. PRÉVOIR AVEC UN MODÈLE AUTORÉGRESSIF VECTORIEL

On admet en modélisation macroéconomique que le modèle autorégressif vectoriel (VAR), proposé par SIMS en 1980, constitue une solution de rechange valable pour le modèle économétrique à équations simultanées standard (SEM). On peut considérer le modèle VAR comme une forme réduite dans laquelle chaque variable endogène est régressée en fonction de ses propres valeurs passées et des valeurs passées des autres variables dans le système. Le modèle VAR peut en outre offrir une méthode de prévision pratique surtout en l'absence d'une théorie a priori solide sur la spécification du modèle.

On a proposé dans la littérature deux extensions du système VAR de Sims. La première, mise de l'avant par Hsiao en 1981, réduit le nombre de paramètres à estimer dans un modèle VAR. Contrairement au VAR de Sims dans lequel toutes les variables du système ont le même retard, le système de Hsiao permet aux variables du système d'avoir des retards différents d'une équation à l'autre. La seconde extension du travail de Sims est l'approche bayésienne de Litterman (1979), dans laquelle de l'information a priori sur les paramètres du modèle est incorporée dans les estimations du modèle.

Dans le présent chapitre nous analyserons les modèles VAR. Nous décrirons tout d'abord les trois types de modèles VAR en les illustrant au moyen de systèmes VAR bivariés que nous avons élaborés traitant des mises en chantier et des consentements hypothécaires au Canada. Nous comparerons ensuite leur efficacité prévisionnelle en analysant leurs prévisions a posteriori.

5.1 LE SYSTÈME VAR DE SIMS

Considérons M séries temporelles liées. On peut décrire une série d'observations conjointes sur ces séries à l'instant t au moyen d'un vecteur M x 1
$x_t = (X_{t1}, X_{t2}, \ldots, X_{tM})$. Les données à disposition consistent en T observations conjointes $x_1, x_2, \ldots, x_T$. On définit comme suit une autorégression vectorielle d'ordre p :

$$x_t = \Phi_1 x_{t-1} + \cdots + \Phi_p x_{t-p} + u_t, \tag{5.1}$$

où

$$\Phi_i = \begin{pmatrix} \phi_{11,i} & \cdots & \phi_{1M,i} \\ \vdots & \ddots & \vdots \\ \phi_{M1,i} & \cdots & \phi_{MM,i} \end{pmatrix}; \qquad i = 1, \ldots, p$$

est une matrice à coefficients M x M. Le terme de perturbation aléatoire u_t est un bruit blanc vectoriel de moyenne zéro et matrice des covariances Σ.

On peut exprimer de manière algébrique une équation représentative d'un système VAR (p) comme suit :

$$X_{tj} = \phi_{j1,1}X_{t-1,1} + \cdots + \phi_{j1,p}X_{t-p,1} + \cdots + \phi_{jM,1}X_{t-1,M} + \cdots + \phi_{jM,p}X_{t-p,M} + u_{tj}, \qquad (5.2)$$

où le nombre de coefficients est Mp. Notons que les M variables sont traitées symétriquement en ce sens que chaque variable est déterminée par ses propres valeurs retardées et par les valeurs retardées des autres variables du système (retards croisés). Le système a donc l'apparence d'un ensemble de régressions linéaires multiples contenant le même ensemble de régresseurs, aussi chaque équation peut-elle être estimée séparément au moyen des moindres carrés conditionnels.

Nous avons construit un modèle VAR bivarié des mises en chantier et des consentements hypothécaires au Canada. Les données utilisées comprennent 252 observations mensuelles des mises en chantier au Canada s'échelonnant de janvier 1968 à décembre 1988 (indice CANSIM D84975) et les traditionnels consentements hypothécaires (indice CANSIM D2653). La série des mises en chantier est la même que celle qui a été utilisée au Chapitre 3 pour la modélisation ARIMA. Les logarithmes népériens des deux séries ont été différenciés consécutivement et saisonnièrement une fois chacun pour obtenir la stationnarité des séries.

Bien qu'il soit facile d'estimer le système VAR de Sims avec les moindres carrés, il n'existe pas de prescription simple permettant de déterminer le nombre de retards nécessaire. Étant donné que les variables ont un caractère saisonnier marqué, p doit être de nombre élevé pour que le modèle estimé reproduise l'effet saisonnier. Toutefois, le nombre de degrés de liberté disponibles pour l'estimation décroît rapidement à mesure que s'accroît l'ordre p. Au moyen du critère AIC d'Akaike (1974), nous avons établi l'ordre à la valeur p = 36. Le nombre total de paramètres à estimer dans une équation est alors de 73, y compris le terme constant.

Les résultats de l'estimation sont présentés au Tableau 5.1[1]. Nous avons déterminé que nos estimations des coefficients

---

[1] Nous avons utilisé le programme RATS, Version 3.1, de VAR Econometrics, pour l'estimation de l'autorégression vectorielle exposée dans ce chapitre.

de retard sont statistiquement significatives pour les intervalles de retard inférieurs et saisonniers, alors que la plupart des estimations des coefficients de retard croisés sont non significatives même pour les retards saisonniers. Mises à part les plus récentes, les valeurs retardées des consentements hypothécaires ne semblent pas être utiles pour expliquer les mises en chantier. Les prévisions tirées du modèle VAR ajusté seront examinées à la Section 5.4.

5.2 LE SYSTÈME VAR DE HSIAO

La plupart des modèles VAR utilisés en macroéconométrie comprennent de nombreuses variables endogènes et de nombreuses valeurs retardées pour chaque variable. Le principal désavantage des modèles VAR est que le nombre de paramètres à estimer s'accroît très rapidement à mesure qu'augmente le nombre de variables et/ou de valeurs retardées. Le perfectionnement par Hsiao (1981) du modèle VAR de Sims réduit l'ampleur du problème relatif aux degrés de liberté en permettant aux intervalles de retard des diverses variables dans chaque équation d'être de longueur différente. En fait, il n'y a <u>a priori</u> aucune raison pour laquelle les variables devraient toutes avoir la même longueur d'intervalle dans toutes les équations du modèle.

Soit X_1 la première variable endogène du système, par exemple la série des mises en chantier évoquée plus haut. Nous inspirant de la stratégie de modélisation de Hsiao, nous avons au départ spécifié l'équation des mises en chantier sous la forme de l'autorégression

$$X_{t1} = \phi_{11}(B) X_{t1} + u_{t1}, \qquad (5.3)$$

où $\phi_{11}(B)$ est un polynôme dans l'opérateur B. Nous avons déterminé l'ordre p de $\phi_{11}(B)$ en minimisant la FPE d'Akaike (1969, 1970) dans un domaine d'ordres possibles allant de 1 à 48 :

$$FPE(p) = \frac{T+p}{T-p} \cdot \frac{SSR(p)}{T},$$

où T = 227 est le nombre d'observations effectives après différenciation et SSR(p) est la somme des termes résiduels au carré du modèle AR(p) ajusté. La FPE minimale = 0,02935 a été observée pour un décalage de 36. Nous avons omis le terme d'intersection dans l'autorégression puisqu'il n'était pas significatif dans le modèle VAR de Sim.

Une fois l'ordre de l'autorégression déterminé, nous avons ajouté la seconde variable du système X_2, les consentements hypothécaires, à l'équation d'autorégression (5.3) pour obtenir la relation bivariée de la forme :

$$X_{t1} = \phi_{11}(B) X_{t1} + \phi_{12}(B) X_{t2} + u_{t1}. \qquad (5.4)$$

L'ordre de $\phi_{12}(B)$ de X_2 a ensuite été déterminé en minimisant la FPE dans un domaine d'ordres possibles allant de 1 à 48, l'ordre de $\phi_{11}(B)$ de X_1 étant maintenu à la valeur antérieurement déterminée de 36. Si la FPE dans la régression bivariée n'avait pas été inférieure à celle de l'autorégression de X_1, alors la variable des consentements hypothécaires X_2 n'aurait pas été utile pour prévoir la variable des mises en chantier X_1 (c'est-

à-dire que X_2 n'agit pas causalement « à la Granger » sur X_1); X_2 aurait donc été supprimée de l'équation en X_1[1]. La FPE minimale = 0,02647 a été observée pour la longueur d'intervalle croisé de 22. Nous avons alors réévalué la longueur optimale des intervalles de retard propres en réestimant (5.4) pour un domaine d'ordres possibles allant de 1 à 48 tout en maintenant la longueur des intervalles croisés à la valeur de 22. Nous avons constaté que, comme auparavant, la FPE minimale s'est maintenue à la même valeur de l'ordre, soit 36, pour les intervalles propres.

L'équation de régression pour les consentements hypothécaires a été déterminée selon un procédé semblable. La FPE minimale = 0,05364 a été observée lorsque l'ordre des intervalles propres était de 36 et celui des intervalles croisés de 8.

Il s'est révélé que le système de Hsiao identifié ci-dessus avait la même longueur d'intervalle propre que le VAR de Sims mais, comme prévu, les intervalles croisés sont beaucoup plus courts dans le premier que dans le second. On trouvera au Tableau 5.2 les résultats de l'estimation du système à deux équations au moyen de la méthode SUR de Zellner. Comme dans le cas du modèle VAR de Sims, les estimations des coefficients des intervalles propres sont significatives pour les intervalles inférieurs et saisonniers. Conséquence de l'application du critère de minimalisation de la FPE, de nombreux intervalles croisés ont été omis au stade de la spécification du modèle. On constate que les erreurs-types des estimations des coefficients dans le modèle VAR de Hsiao soit beaucoup plus petites que celles du modèle VAR de Sims. Le surajustement et le sous-ajustement des intervalles du système indiquent que le modèle ajusté est adéquat.

5.3 LE SYSTÈME VAR BAYÉSIEN

Dans l'approche bayésienne de la modélisation VAR, les prévisionnistes spécifient un modèle comme en (5.1) avec un intervalle long abordable par le calcul et imposent des conditions bayésiennes d'une forme ou l'autre à tous les coefficients du système. Nous avons élaboré deux modèles VAR bayésiens des mises en chantier et des consentements hypothécaires, modèles que nous présentons dans la présente section. On trouvera des analyses des développements récents en

---

[1] Nous avons renoncé à élaborer un modèle VAR des mises en chantier et des taux hypothécaires parce qu'on a constaté que les taux hypothécaires n'agissaient pas causalement « à la Granger » sur les mises en chantier.

modélisation VAR bayésienne dans Todd (1984), Doan, Litterman et Sims (1984) et Litterman (1986b).

Dans le premier modèle VAR bayésien des mises en chantier et des consentements hypothécaires nous faisons emploi des mêmes variables dépendantes que dans les systèmes de Sims et de Hsiao, c'est-à-dire de la série différenciée stationnaire. Puisque les variables sont exprimées en différences, nous avons introduit les conditions dans le cadre de la spécification d'un processus de bruit blanc avec dérive :

$$X_{it} = \delta + u_{it}, i = 1,2. \tag{5.5}$$

On sait très peu de choses a priori à propos des paramètres de dérive, aussi utilisons-nous dans ce contexte des a priori non informatifs. Compte tenu de (5.1) la spécification de bruit blanc en (5.5) implique que tous les coefficients de retard sont de valeur nulle. Nous employons donc pour les coefficients de retard de chaque équation des conditions normales indépendantes de moyenne nulle. Leurs déviations standard a priori ne sont pas nulles dans la spécification.

L'utilisation des déviations standard a priori de type courant se fonde sur l'idée que les valeurs plus récentes d'une variable sont plus susceptibles de renfermer des informations utiles pour la prévision que les valeurs moins récentes. C'est pourquoi les déviations standard a priori sont spécifiées avec les caractéristiques suivantes :

(1) elles diminuent à mesure que s'accroissent les retards; et
(2) elles sont plus grandes pour les coefficients de retard propres que pour les coefficients de retard croisés des autres variables du système.

On peut spécifier les déviations standard a priori dotées de telles caractéristiques en termes de quelques « hyperparamètres » :

$$s(i,j,h) = \gamma g(h) f(i,j) (s_i/s_j). \tag{5.6}$$

où s(i, j, h) est la déviation standard à priori pour le $h^{ième}$ coefficient de retard de la $j^{ième}$ variable de la $i^{ième}$ équation; s_i et s_j sont les erreurs-types de l'autorégression des variables i et j respectivement; g(h) est le paramètre de désintégration pour la cohésion de l'intervalle de retard h relativement à celle du premier intervalle (g(1) = 1); et γ est la cohésion globale, qui est égale à la déviation standard du premier intervalle propre.

Les déviations standard a priori du type de base décrit ci-dessus s'appliquent aux variables non stationnaires ou qui ont été saisonnièrement ajustées. Étant donné que la désintégration affectant la déviation standard lorsque s'accroît le retard peut rendre imprécises les estimations des paramètres pour les retards saisonniers importants, nous avons décidé d'établir g(h) = 1 pour tous les intervalles de retard, c'est-à-dire que les déviations standard a priori ne se désintègrent pas avec l'accroissement du retard. Les déviations standard sont alors

$$s(i,j,h) = \gamma f(i,j)(s_i/s_j) \qquad (5.7)$$

et

$$f(i,j) = \begin{cases} 1 & \text{si } i = j \\ \omega & \text{dans les autres cas.} \end{cases}$$

Étant donné les valeurs des deux hyperparamètres γ et , on combine les conditions et les données pour estimer un modèle VAR de (5.1) au moyen de la méthode d'estimation mixte de Theil ou de la procédure de régression (SUR) apparemment indépendante de Zellner. Bien que la méthode SUR procure un gain d'efficacité, nous avons utilisé celle de Theil pour des raison de facilité de calcul.

Les hyperparamètres ne sont pas spécifiés a priori mais choisis empiriquement par recherche dans le domaine de leurs valeurs possibles et en utilisant un critère prévisionnel. Étant donné une paire de valeurs pour γ et ω, on estime les paramètres du modèle et on utilise le VAR ajusté pour faire des prévisions post-échantillonnage. Nous avons choisi comme critère de sélection les erreurs absolues moyennes en pourcentage (MAPE) ou la moyenne quadratique des erreurs en pourcentage $(RMSPE)_{10}$ des prévisions à une à dix périodes vers l'avenir[3].

Après une recherche limitée parmi les valeurs possibles de γ et ω, nous avons obtenu les valeurs minimales de MAPE = 7,22 et RNSPE = 8,61 pour γ = 0,5 et ω = 0,2. L'utilisation d'une grille

---

[3] Nous avons fait usage du critère choisi à cause de contraintes de temps et de calcul. Il est possible de calculer les MAPE ou RMSPE pour chaque variable en effectuant plusieurs prévisions pour chaque horizon prévisionnel et de choisir les valeurs des hyperparamètres à partir des meilleurs résultats globaux. Doan, Litterman et Sims (1984) ont proposé l'emploi de méthodes de recherche des valeurs des hyperparamètres qui minimisent la somme des RMSPE à une étape vers l'avenir.

de recherche plus fine des paramètres γ et ω aurait rapporté des valeurs plus petites de la MAPE et de la RMSPE, mais nous n'avons pas tenté de ce faire faute de temps. Nous avons toutefois noté que l'efficacité prévisionnelle n'est par particulièrement sensible aux variations du paramètre ω.

Le Tableau 5.3 décrit le modèle VAR bayésien ajusté avec les valeurs des hyperparamètres telles que données ci-dessus. Notons que la plupart des estimations de coefficients ne sont pas statistiquement significatives relativement aux moyennes a priori égales à zéro, exception faite de quelques-unes pour les premiers retards propres ou retards saisonniers propres. Les erreurs-types des estimations des coefficients s'avèrent plus petites que celles dans le modèle VAR de Hsiao ajusté.

Nous avons également mis à l'essai une autre formulation VAR bayésienne. Rappelons-nous que dans les trois modèles VAR précédents nous avons transformé les données pour obtenir la stationnarité. Or une telle transformation peut ne pas être nécessaire dans un modèle VAR. Fuller (1976, p. 374) démontre que la différenciation ne procure aucun gain d'efficacité asymptotique pour l'autorégression même lorsque la série est intégrée. De plus, lorsque des séries diffèrent quant à leur ordre d'intégration, la différenciation ne fournit pas un moyen satisfaisant pour élaborer des modèles VAR de séries temporelles non stationnaires (Harvey (1989), p. 469). C'est pourquoi nous avons également conçu un modèle VAR bayésien de la série à paliers et utilisé les a priori relatifs aux paramètres du modèle en nous fondant sur le modèle de la marche aléatoire avec dérive :

$$X_{ti} = \delta + X_{t-1,i} + u_{ti}. \qquad (5.8)$$

Nous avons donné la valeur de un à la moyenne a priori du coefficient du premier intervalle de retard propre. Les moyennes a priori de tous les autres paramètres ont été fixées à zéro.

Nous avons traité le caractère saisonnier de la série de manière déterministe et avons inclus 12 variables fictives mensuelles dans le système. Il s'agit là d'une approche préconisée par Doan (1990). Nous avons fait usage d'a priori non informatifs pour les coefficients des variables fictives saisonnières. Nous avons établi les déviations standard a priori des autres paramètres de manière à obtenir une désintégration harmonique normale, de sorte que g(h) = 1/h en (5.5)[4].

---

[4] Raynauld et Simonato (1988, 1990) ont également tenté de modifier les caractéristique de la méthode VAR bayésienne de base pour tenir compte des composantes saisonnières de la série temporelle. Nous avons mis à l'essai leur procédé à désintégration « en dents de scie », mais le modèle VAR ajusté n'a pas fourni des prévisions aussi précises que celui ajusté au moyen d'une

Le modèle ajusté final choisi après une recherche limitée parmi les valeurs possibles des hyperparamètres γ et ω est décrit au Tableau 5.4. Les valeurs de $\gamma = 2,5$ et $\omega = 0,3$ ont fourni les valeurs minimales de la MAPE = 7,06 et la RMSPE = 8,83. Il est intéressant de noter que dans les deux équations les coefficients des retards croisés et des retards propres sont significatifs pour les premiers et/ou seconds intervalles de retard. Les consentements hypothécaires contribuent en fait à expliquer les variations du nombre des mises en chantier. Aucune des estimations des coefficients pour les intervalles saisonniers n'est significative. Il semble enfin que les variables fictives saisonnières ont rendu compte adéquatement des variations saisonnières de la série.

5.4 EFFICACITÉ PRÉVISIONNELLE DES MODÈLES VAR

Lorsqu'on dispose d'un modèle VAR, on peut l'utiliser pour obtenir des prévisions conjointes de toutes les variables du système. Comme dans le cas des modèles ARIMA univarié et à fonction de transfert, l'espérance de x_{T+h} dépendante de l'historique I_T du processus vectoriel jusqu'à l'instant T correspond à la prévision à h périodes vers l'avenir optimale quant à la moyenne quadratique minimale des erreurs :

$$f_T(h) = E_T[x_{T+h}|I_T],$$

où $f_T(h)$ est la prévision à h étapes vers l'avenir optimale et E_T l'opérateur d'espérance conditionnelle, tous deux à l'instant T. On peut calculer par récurrence les prévision à h étapes vers l'avenir à partir de

$$\begin{aligned} f_T(h) &= E_T(x_{T+h}) \\ &= \Phi_1 f_T(h-1) + \cdots + \Phi_p f_T(h-p), \quad h = 1,2,\ldots \end{aligned} \tag{5.9}$$

où $f_T(h-i) = x_{T+h-i}$ for $i \geq h$.

Si les erreurs aléatoires sont un processus de bruit blanc normal, les erreurs de prévision sont également normales :

---

désintégration simple. Nous avons en outre mis à l'essai un modèle sans désintégration des déviations standard. Contrairement à ce qui est le cas avec un modèle VAR exprimé en différences, la capacité prévisionnelle a été inférieure à celle du modèle VAR bayésien avec désintégration déjà évoqué.

$$x_{T+h} - f_T(h) \sim N[0, \Sigma(h)], \qquad (5.10)$$

où $\sum(h)$ est la matrice des covariances des erreurs prévisionnelles, qui dépend notamment des Φ. On peut donc aménager une région de confiance pour un ensemble de prévisions ou pour une prévision unique ponctuelle. On peut par exemple obtenir à partir de (5.10) une prévision avec un intervalle de confiance de 95 p. 100 pour une composante unique de x_{T+h} sous la forme

$$f_T^m(h) \pm 1.96\sigma_m(h),$$

où $f_T^m(h)$ est la $m^{ième}$ composante de $f_T(h)$ et $\sigma(h)$ est la déviation standard de la prévision ponctuelle égale à la racine carrée du $m^{ième}$ élément diagonal de $\sum(h)$.

En pratique, les paramètres du modèle VAR sont inconnus et estimés. On peut remplacer les paramètres inconnus dans les expressions prévisionnelles, les intervalles de prévision ainsi obtenus n'étant que des intervalles approximatifs de 95 p. 100.

Pour évaluer l'efficacité prévisionnelle des modèles VAR, nous avons effectué des prévisions ARIMA des mises en chantier en 1989. Le modèle ARIMA ajusté au moyen des données échelonnées de janvier 1968 à décembre 1988 est semblable au « Modèle 2 du Chapitre 3, sauf que le terme AR(1) a été omis :

$$(1 + .2383B^{12} + .1227B^{24} + .2557B^{36})(1 - B)(1 - B^{12})\ LOGHS_t$$

$$= (1 - .2043B)(1 - .5461B^{12})\underline{u}_t.$$

Toutes les estimations de coefficients sont significatives sauf l'estimation du paramètre saisonnier AR(2). Diverses vérifications diagnostiques indiquent que le modèle ajusté est adéquat.

Le Tableau 5.5 rassemble les prévisions de un à dix mois vers l'avenir des mises en chantier (par paliers) fondées sur les quatre modèles VAR de même que celles fondées sur le modèle ARIMA ci-dessus. Les erreurs prévisionnelles effectives du modèle ARIMA varient en valeur absolue de 0,680 pour la prévision dix mois vers l'avenir d'octobre 1989 à 11,35 p. 100 pour la prévision huit mois vers l'avenir d'août 1989. Si la prévision un mois vers l'avenir est mauvaise, puisque inexacte de 11 p. 100, les prévisions ARIMA ne paraissent pas se dégrader à mesure que s'éloigne l'horizon prévisionnel.

Le modèle VAR de Sims a fourni des prévisions très précises pour un horizon à court terme. Après la prévision trois mois vers l'avenir, sa capacité prévisionnelle s'est détériorée rapidement, ses erreurs prévisionnelles variant en valeur absolue

de 9,7 p. 100 à 32,47 p. 100 par rapport aux mises en chantier effectives. S'agissant des erreurs prévisionnelles individuelles, de la MAPE ou de la RMSPE, il n'a pas fait mieux que le simple modèle ARIMA univarié.

Le modèle VAR de Sims est facile à spécifier et à estimer. Le principal désavantage qu'il présente en tant qu'outil prévisionnel est lié au problème du nombre de degrés de liberté. En effet, lorsque le modèle comprend de nombreux paramètres à estimer, la perte résultante d'efficacité dans l'estimation des paramètres donne lieu à des erreurs standard élevées dans les prévisions. Cette perte d'efficacité explique la mauvaise prestation du modèle VAR de Sims pour la prévision des mises en chantier en 1989.

Le modèle VAR de Hsiao ajusté a fourni de meilleures prévisions que le modèle VAR de Sims. Pour l'horizon à court terme jusqu'à trois mois, ses prévisions sont très précises. Quoique ses prévisions se dégradent à mesure que s'éloigne l'horizon prévisionnel, ses erreurs prévisionnelles sont systématiquement beaucoup plus réduites que celles du modèle VAR de Sims non contraint pour tous les horizons prévisionnels de quatre à dix mois.

Comme nous l'avons mentionné dans la section précédente, l'élaboration d'un modèle VAR de Hsiao est plus coûteuse en calculs qu'un modèle VAR de Sims, car le premier traite explicitement la longueur d'intervalle pour chaque variable de chaque équation. Mais l'approche de Hsiao réduit l'acuité du problème du nombre des degrés de liberté en modélisation VAR et fournit des estimations des paramètres plus efficaces et des prévisions plus précises que le modèle VAR de Sims. Mais en ce qui a trait à la MAPE et à la RMSPE, toutefois, il ne présent aucun avantage par rapport au modèle ARIMA univarié pour la prévision des mises en chantier.

Nous l'avons vu au Chapitre 2, des travaux considérables ont été accomplis pour comparer les prévisions macroéconomiques a priori des modèles économétriques standard (SEM) à grande échelle avec celles des modèles ARIMA. Récemment, Litterman (1986) et McNees (1986, 1988) ont utilisé des modèles VAR bayésiens à des fins de comparaison. Litterman a constaté qu'en moyenne les prévisions fournies par son modeste modèle VAR bayésien de l'économie américaine sont aussi précises que celles obtenues par les services de prévision commerciaux éminents au moyen de SEM complexes à grande échelle. De même, McNees a observé qu'un petit modèle VAR bayésien de l'économie américaine prédit certaines variables très bien, même pour des horizons temporels éloignés.

L'efficacité des modèles VAR bayésiens des mises en chantier que nous avons élaborés se compare favorablement à celle des deux autres modèles VAR. Sous leur forme différenciée, ils ont fourni des prévisions dont la valeur absolue des erreurs varie de 0,95 p. 100 pour une prévision quatre mois vers l'avenir (avril 1989) à 12,17 p. 100 pour une prévision un mois vers l'avenir (janvier 1989). Mises à part les prévisions un et deux mois vers l'avenir, le modèle VAR bayésien a engendré des erreurs prévisionnelles qui sont inférieures en valeur absolue à celles engendrées par le modèle VAR de Hsiao. De plus, la MAPE et la RMSPE sont de beaucoup inférieures dans le modèle VAR bayésien à ce qu'elles sont dans le modèle VAR de Hsiao.

Le modèle VAR bayésien sous sa forme à paliers est aussi efficace que sous sa forme différenciée. Mise à part la prévision un mois vers l'avenir pour janvier 1989, le modèle VAR bayésien à paliers semble avoir fait meilleure figure quant à l'amplitude des erreurs prévisionnelles individuelles de même qu'à la MAPE et la RMSPE.

Un avantage important que présentent les modèles VAR par rapport aux SEM traditionnels est le fait qu'on peut estimer un modèle VAR au moyen de données remontant à tout moment du passé, fournir des prévisions pour un horizon aussi éloigné que voulu et tester la précision prévisionnelle du modèle. C'est exactement cela que nous avons fait dans ce chapitre.

Comparés au modèle économétrique traditionnel, tous les types de modèles VAR examinés plus haut présentent les avantages suivants en tant qu'instruments de prévision :

(1) le modèle VAR ne requiert pas d'ajustement subjectif de ses paramètres ou de ses prévision;
(2) il n'opère pas la douteuse distinction d'exogénéité parmi les variables à prédire;
(3) il offre une méthode conceptuellement claire et simple pour fournir des prévisions non conditionnelles qui n'ont pas besoin de prendre en compte les valeurs futures de variables exogènes; et
(4) il n'impose pas les restrictions théoriques controversées qui peuvent caractériser le SEM traditionnel.

Le modèle VAR bayésien a fait meilleure figure que le modèle ARIMA. Son principal point faible comparé au modèle ARIMA ou au modèle VAR de Hsiao est le fait qu'il est de loin le plus coûteux des modèles à élaborer. En outre, le traitement efficace de la saisonnalité dans les séries temporelles en modélisation VAR bayésienne en est encore à un stade exploratoire dans la littérature actuelle sur les séries temporelles. Lorsqu'on aura maîtrisé le traitement de la saisonnalité en modélisation VAR, il

est à prévoir que les prévisions VAR bayésiennes concurrenceront efficacement les prévisions ARIMA.

Tableau 5.1

RÉSUMÉ D'UN MODÈLE VAR DE SIMS DES MISES EN CHANTIER ET DES CONSENTEMENTS HYPOTHÉCAIRES

	(1) VARIABLES DE L'ÉQUATION DES MISES EN CHANTIER		(2) VARIABLES DE L'ÉQUATION DES CONSENTEMENTS	
RETARD	MISES EN CHANTIER	CONSENTEMENTS	MISES EN CHANTIER	CONSENTEMENTS
1	-.343*	.113*	.232*	-.101
2	-.318*	.173*	-.026	-.154
3	-.280*	.176*	-.050	-.091
4	-.200*	.121*	-.192	-.019
5	-.146	.109	-.115	-.032
6	-.090	.051	-.257	-.034
7	-.066	.040	-.145	.058
8	.007	.011	-.076	-.021
9	-.137	.054	-.005	.137
10	.038	-.115	-.111	.017
11	-.029	.083	-.069	.043
12	-.689*	-.034	-.058	-.581*
13	-.317*	.028	-.043	-.030
14	-.145	.089	.049	-.245*
15	-.174	.104	.030	-.014
16	-.185	.036	-.121	-.093
17	-.036	.066	-.018	-.133
18	-.107	.077	-.043	-.078
19	-.040	-.080	-.016	-.127
20	.141	-.110	-.067	-.214*
21	-.046	-.054	-.153	.021
22	.054	-.110	-.153	-.023
23	-.005	.028	-.149	.181
24	-.289*	-.032	-.130	-.286*
25	-.159	.039	-.051	.005
26	-.031	.023	.020	-.085
27	-.153	.011	.015	.061
28	-.114	-.017	-.024	.034
29	-.045	.024	-.059	.007
30	-.042	.062	.102	-.065
31	-.104	-.027	.077	-.146
32	.036	.019	.112	-.181
33	-.051	.003	.002	-.105
34	-.129	.017	.012	-.097
35	.007	.045	.046	.159
36	-.321*	.011	-.150	-.258*
CONSTANTE	.002		-.001	
	R**2=.634		R**2=.534	
	SEE =.1539		SEE =.2256	

NOTE : L'astérisque * indique que l'estimation du coefficient est significative au seuil de signification = 0,05.

Tableau 5.2

RÉSUMÉ D'UN MODÈLE VAR DE HSIAO DES MISES EN CHANTIER ET DES CONSENTEMENTS HYPOTHÉCAIRES

	(1) VARIABLES DE L'ÉQUATION DES MISES EN CHANTIER		(2) VARIABLES DE L'ÉQUATION DES CONSENTEMENTS	
RETARD	MISES EN CHANTIER	CONSENTEMENTS	MISES EN CHANTIER	CONSENTEMENTS
1	-.388*	.102*	.172*	-.147*
2	-.329*	.181*	-.089	-.150*
3	-.246*	.148*	-.116	-.128
4	-.127	.088	-.169	-.042
5	-.096	.051	-.194*	-.002
6	-.090	-.011	-.323*	-.066
7	-.089	.041	-.214*	.056
8	-.064	.042	-.187*	.032
9	-.178*	.084		.115
10	-.015	-.084		.023
11	-.004	.064		.052
12	-.683*	-.006		-.598*
13	-.332*	-.013		-.088
14	-.125	.085		-.209*
15	-.102	.075		-.027
16	-.085	.013		-.112
17	.004	.037		-.118
18	-.083	.012		-.155*
19	-.060	-.085		-.128
20	.119	-.129*		-.198*
21	-.011	-.074		-.052
22	.049	-.129*		-.060
23	.055			.144*
24	-.254*			-.346*
25	-.133*			-.057
26	-.030			-.073
27	-.119*			.027
28	-.096			.048
29	-.022			-.044
30	-.015			-.036
31	-.085			-.064
32	-.073			-.099
33	-.018			-.066
34	-.129*			-.046
35	-.003			.185*
36	-.294*			-.257*
	R**2=.652		R**2=.518	
	SEE =.1457		SEE =.2122	

Note : L'astérisque * indique que l'estimation du coefficient est significative au seuil de signification = 0,05.

Tableau 5.3

RÉSUMÉ D'UN MODÈLE VAR BAYÉSIEN DES MISES EN CHANTIER ET DES CONSENTEMENTS HYPOTHÉCAIRES (EN DIFFÉRENCES)

	(1) VARIABLES DE L'ÉQUATION DES MISES EN CHANTIER		(2) VARIABLES DE L'ÉQUATION DES CONSENTEMENTS	
RETARD	MISES EN CHANTIER	CONSENTEMENTS	MISES EN CHANTIER	CONSENTEMENTS
1	-.081*	.001	.001	-.029
2	-.041	.005	-.003	-.031
3	-.018	.001	-.002	-.042
4	-.021	-.000	-.002	-.009
5	-.043	.001	.002	-.024
6	.007	-.001	-.006	-.029
7	.009	.002	-.003	.024
8	-.002	.002	-.001	-.000
9	-.027	.002	.005	.048
10	.037	-.003	-.002	.012
11	-.030	.002	.001	-.015
12	-.275*	-.004	-.002	-.238*
13	-.047	-.002	-.010	-.001
14	.022	-.001	.003	-.056
15	.003	.001	.000	.024
16	-.009	-.001	-.000	.005
17	.018	.001	-.000	-.034
18	-.029	.001	.004	.019
19	-.010	-.003	.001	-.004
20	.056	-.004	.006	-.048
21	-.002	.000	-.005	-.020
22	.016	.001	-.003	.024
23	.029	.001	-.002	.075
24	-.023	.002	.002	-.057
25	-.007	.003	.003	.013
26	.011	-.001	-.001	.032
27	-.036	-.003	.001	.008
28	-.002	-.001	-.000	.003
29	.012	-.000	-.005	.009
30	.022	.002	.002	.003
31	-.028	.001	.000	-.029
32	.026	.004	-.002	-.007
33	-.008	-.001	.002	.006
34	-.073	.001	.000	-.019
35	.033	-.002	.005	.063
36	-.134*	-.002	-.007	-.097*
CONSTANTE	-.002		-.002	
	R**2=.355		R**2=.296	
	SEE =.1705		SEE =.2316	

NOTE : L'astérisque * indique que l'estimation du coefficient est significative au seuil de signification = 0,05.

Tableau 5.4

RÉSUMÉ D'UN MODÈLE VAR BAYÉSIEN DES MISES EN CHANTIER
ET DES CONSENTEMENTS HYPOTHÉCAIRES
AVEC VARIABLES FICTIVES SAISONNIÈRES
(PAR PALIERS)

	(1) VARIABLES DE L'ÉQUATION DES MISES EN CHANTIER		(2) VARIABLES DE L'ÉQUATION DES CONSENTEMENTS	
RETARD	MISES EN CHANTIER	CONSENTEMENTS	MISES EN CHANTIER	CONSENTEMENTS
1	.619*	.086	.213*	.825*
2	.011	.144*	-.225*	-.011
3	-.025	-.064	.047	.010
4	.088	-.020	-.088	.038
5	-.041	.071	.039	-.023
6	.030	-.064	-.063	-.009
7	.043	-.004	.081	.082
8	.075	-.012	.023	-.050
9	-.065	-.006	.021	.093
10	.070	-.056	-.000	.079
11	.031	.027	-.004	.059
12	-.001	-.044	-.013	-.015
13	-.032	-.008	-.017	-.012
14	.044	.003	.062	-.125
15	-.017	-.000	.015	.151*
16	-.000	.001	.013	-.059
17	-.053	.011	.026	-.057
18	-.085	-.009	.040	.013
19	-.011	-.025	.002	-.036
20	.090	-.003	-.018	-.059
21	-.075	.023	-.055	.012
22	-.023	.015	-.004	.072
23	.092	.009	.001	.114
24	.035	-.002	-.003	-.082
25	-.012	.001	.006	.020
26	-.055	-.008	-.005	-.004
27	-.032	-.002	-.011	.027
28	.021	.009	-.006	.021
29	.014	.010	-.002	-.011
30	-.015	.010	.020	-.009
31	-.048	.009	.006	-.042
32	-.018	.012	-.005	-.026
33	-.022	.004	.003	-.031
34	-.055	.000	.008	.002
35	.082	-.010	.007	.046
36	-.003	-.006	-.009	-.031
	R**2=.880		R**2=.848	
	SEE =.1246		SEE =.1787	

Note : L'astérisque * indique que l'estimation du coefficient est significative au seuil de signification = 0,05. Toutes les 12 variables fictives étaient significatives dans l'équation des mises en chantier et aucune n'était significative dans l'équation des consentements.

Tableau 5.5

ERREURS DE PRÉVISION SUR UNE À DIX PÉRIODES VERS L'AVENIR
DES MISES EN CHANTIER FONDÉES SUR CINQ MODÈLES

MOIS 1989	VALEUR EFFECTIVE	VAR DE SIMS	VAR DE HSIAO	VAR BAYÉSIEN (DIFFÉRENCES)	VAR BAYÉSIEN (PALIERS)	ARIMA
Jan.	13,678	794 (5.81)	1056 (7.72)	1,664 (12.17)	2,506 (18.32)	1,504 (11.00)
Fév.	9,774	103 (1.05)	−90 (0.92)	845 (8.64)	1,087 (11.12)	459 (4.69)
Mar.	12,150	−673 (−5.55)	−847 (−6.97)	324 (2.67)	1,227 (10.09)	1,066 (8.77)
Avr.	17,392	−1,693 (−9.74)	−1,179 (−6.78)	164 (0.95)	1,643 (9.44)	816 (4.70)
Mai	19,127	−6,210 (−32.47)	−3,888 (−20.33)	−2,197 (11.49)	814 (4.25)	−1,869 (−9.77)
Juin	17,697	−3,926 (−22.19)	−2,894 (−16.36)	−1,106 (−6.25)	−477 (−2.69)	−1,863 (−10.53)
Juil.	16,421	−3,773 (−22.98)	−3,330 (−20.64)	−821 (−5.00)	−204 (−1.24)	−661 (−4.03)
Août	1,543	−4,635 (−29.94)	−3,350 (−21.63)	−2,587 (−16.71)	−22 (−0.14)	−1,757 (−11.35)
Sep.	16,338	−2,879 (−17.62)	−963 (−5.90)	810 (4.96)	1,526 (9.34)	301 (1.84)
OCT	15,764	−3,244 (−20.38)	−1,268 (−8.04)	532 (3.38)	622 (3.45)	−108 (−0.68)
MAPE		16.79	11.53	7.22	7.06	6.73
RMSPE		19.66	13.53	8.61	8.83	7.72

Note : Les chiffres entre parenthèses sont les erreurs prévisionnelles en pourcentage par rapport aux valeurs effectives.

6. CONCLUSIONS

L'objectif principal de la présente étude consiste à passer en revue diverses méthodes de prévision fondées sur un modèle susceptibles d'être utilisées pour prévoir l'évolution du secteur du logement et des marchés hypothécaires. Les modèles considérés sont (1) le modèle économétrique à équations simultanées (SEM), (2) le modèle autorégressif intégré univarié à moyenne mobile (ARIMA), (3) le modèle à fonction de transfert (TF) et (4) le modèle autorégressif vectoriel (VAR).

Quel que soit le modèle de prévision utilisé, il faut dans tous les cas considérer les questions suivantes pour son élaboration et son utilisation :

(1) spécification pour le modèle d'une équation ou d'équations selon les données et, dans le cas d'un modèle économétrique, les théories économiques pertinentes;
(2) disponibilité et collecte des données;
(3) estimation et évaluation initiales du modèle; et
(4) validation du modèle considéré - vérification diagnostique, prévision a posteriori, simulation et tests de suivi du modèle.

Il faut aborder ces questions avec beaucoup de soin si l'on veut élaborer et exploiter un modèle de prévision fiable.

Jusqu'aux années 1970, le modèle économétrique à équations simultanées était utilisé très couramment comme outil prévisionnel macroéconomique. La spécification d'un modèle économétrique nécessite une représentation explicite des relations d'interdépendance et de causalité supposées entre les variables du modèle. C'est au cours de l'étape de la spécification du modèle que la théorie économique joue un rôle en modélisation économétrique.

En pratique, la prévision avec un modèle économétrique n'est pas fondée uniquement sur le modèle et les données. Un prévisionniste faisant usage d'un modèle économétrique procède à des ajustements subjectifs des paramètres du modèle de même que des prévisions du modèle. C'est à cause de ces ajustements subjectifs que (1) les prévisions fondées sur un modèle économétrique ne peuvent être reproduites sinon par le prévisionniste lui-même et que (2) les erreurs prévisionnelles standard en tant que mesure de la fiabilité des prévisions ne peuvent être calculées pour les prévisions ponctuelles.

De nombreuses raisons expliquent pourquoi l'approche économétrique de la prévision peut fournir des prévisions qui ne sont pas fiables :

(1) il faut prédire les variables exogènes pour prédire les variables endogènes. L'imprécision dans la prévision des variables exogènes entraîne l'imprécision des prévisions conditionnelles de variables endogènes;
(2) certaines variables pourraient à tort être considérées comme exogènes;
(3) s'agissant de variables inobservables comme le revenu fixe et l'inflation anticipée, on fait souvent usage dans le modèle d'ersatz grossiers; et
(4) la théorie économique fournit rarement au concepteur de modèle un guide quant à la spécification dynamique; on introduit souvent les variables endogènes retardées dans les équations structurelles de manière ad hoc.

Au contraire de l'approche économétrique structurelle, les modèles ARIMA et autres modèles à série temporelle ne font qu'un emploi limité de la théorie économique pour la spécification du modèle. Mais même sans ajustements subjectifs, les prévisions ARIMA se sont révélées aussi précises que les prévisions ajustées subjectivement fondées sur des modèles économétriques complexes. En outre, contrairement à ce qui est le cas avec les prévisions fondées sur un modèle économétrique, on peut facilement calculer les erreurs-types des prévisions ARIMA, de sorte qu'une mesure de la précision est associée aux prévisions ponctuelles.

Comme instrument de prévision, le modèle ARIMA univarié présente, comme les autres modèles à série temporelle, les avantages suivants :

(1) il fournit des prévisions très précises, tout au moins à court terme;
(2) il est possible d'estimer un modèle ARIMA en utilisant des données remontant à tout moment dans le passé et d'effectuer ensuite des prévisions permettant de vérifier la précision prévisionnelle du modèle; et
(3) il n'impose aucune des restrictions théoriques controversées que peut contenir le modèle économétrique traditionnel.

Les prévisions tirées d'un modèle ARIMA portent sur une série unique sans faire usage d'informations contenues dans d'autres séries liées. Dans de nombreuses situations prévisionnelles et dans lesquelles d'autres variables influencent la série à prédire, on peut élaborer un modèle à fonction de transfert qui incorpore plus d'une série temporelle et introduit explicitement les caractéristiques dynamiques des séries.

Les prévisions fondées sur un modèle à fonction de transfert peuvent être plus précises que celles tirées d'un modèle ARIMA univarié si la variable d'entrée explique une proportion importante des variations de la variable de sortie. Notons que, tout comme la prédiction des variables endogènes d'un SEM exige

la prédiction des variables exogènes, la prédiction de la variable de sortie d'un modèle à fonction de transfert est tributaire de la prédiction des variables d'entrée. La précision des prévisions portant sur la variable de sortie est donc influencée par celle des prévisions portant sur les variables d'entrée.

Tout comme les modèles ARIMA et à fonction de transfert, le modèle VAR se fonde sur la régularité des variations de séries temporelles sans faire appel à la théorie économique. Les difficultés éprouvées en modélisation économétrique traditionnelle des marchés domiciliaire et hypothécaire indiquent que, en dépit de, ou peut-être grâce à leur approche non théorique, les méthodes de prévision ARIMA et VAR constituent des concurrents prometteurs pour la prévision économétrique.

En macroéconométrie, la plupart des modèles VAR du type de celui de Sims font usage de plusieurs variables et comprennent plusieurs valeurs retardées pour chaque variable. Le nombre de paramètres à estimer dans un modèle VAR s'accroît très rapidement à mesure qu'on augmente le nombre de variables et/ou l'ordre du modèle. Le perfectionnement par Hsiao du modèle VAR de Sims ou la procédure bayésienne de modélisation VAR devraient résoudre partiellement le problème relatif au nombre de paramètres en modélisation VAR. Le modèle VAR bayésien semble être plus efficace à des fins prévisionnelles que les modèles VAR de type Sims ou Hsiao.

En tant que rivale de la méthode de prévision économétrique traditionnelle, la méthode VAR offre les avantages suivants :

(1) elle fournit des prévisions très précises, tout au moins à court terme et peut-être à moyen terme;
(2) contrairement au SEM et au modèle à fonction de transfert, elle ne fait pas usage d'une douteuse définition de l'exogénéité; et
(3) elle constitue une méthode conceptuellement claire et remarquablement simple pour obtenir des prévisions qui ne supposent pas des valeurs particulières pour des variables exogènes.

Lawrence Klein (1982) est d'avis que « les modèles VAR font l'affaire pour les prévisions sur le trimestre suivant, mais les prévisions VAR se détériorent rapidement (avec l'éloignement de l'horizon prévisionnel), de sorte que les méthodes traditionnelles fournissent des prévisions supérieures à plus long terme ». Mais cette constatation ne s'applique pas nécessairement à toutes les situations prévisionnelles. Car les résultats de prévisions présentés dans ce rapport indiquent manifestement que les modèles à série temporelle prévoient les mises en chantier remarquablement bien non seulement jusqu'à un horizon de trois mois mais au moins jusqu'à un horizon de douze

mois. Il est possible que les mises en chantier dans plusieurs mois soient moins dépendantes de la série observée que celles dans un avenir plus immédiat. Aussi la capacité prévisionnelle d'un modèle ARIMA pourra se détériorer assez rapidement à mesure que s'éloigne l'horizon prévisionnel.

On sait peu de choses sur l'efficacité relative des diverses méthodes pour les prévisions à long terme. À mesure que s'éloigne l'horizon prévisionnel, les prévisions fondées sur les modèles ARIMA, VAR ou économétriques tendent vers la moyenne de la série chronologique considérée. Si les modèles économétriques semblent faire meilleure figure que les modèles à série temporelle pour les prévisions à long terme, c'est uniquement parce que le prévisionniste procède à des ajustements subjectifs des valeurs futures anticipées des variables exogènes et des prévisions mécaniques du modèle.

Lorsque s'offrent plusieurs solutions, il faut choisir une méthode prévisionnelle appropriée pour la situation spécifique considérée. Nous avons présenté au Tableau 6.1 un classement des quatre méthodes de prévision selon divers critères prévisionnels importants. Certains critères sont de nature statistique et d'autres de nature pratique : exigences en matière de données, facilité de la spécification du modèle, facilité de l'estimation du modèle, temps requis pour la modélisation, coût de la modélisation, facilité à calculer les prévisions, calcul des erreurs-types des prévisions, incorporation d'ajustements subjectifs des prévisions, précision des prévisions à court terme, précision des prévisions à long terme et facilité de remise à jour du modèle.

Cette liste de critères n'est pas exhaustive, et l'on pourra y ajouter d'autre critères. En outre, ces critères ne sont pas nécessairement d'égale importance pour évaluer l'efficacité des méthodes de prévision. Nonobstant les limites inhérentes à une simple moyenne ou somme, lorsqu'on calcule la somme des rangs respectifs pour les critères énumérés, on obtient le classement suivant (en ordre descendant) : ARIMA, TF, SEM et VAR. Il n'existe pas de meilleure approche ou méthode. Ce qui importe, c'est de comprendre comment les diverses méthodes de prévision diffèrent l'une de l'autre de manière à ce que le prévisionniste puisse opérer un choix rationnel selon ses besoins.

La méthode VAR bayésienne n'est pas encore suffisamment développée pour traiter adéquatement de la saisonnalité en modélisation et en prévision. Par ailleurs, l'approche par modélisation avec fonction de transfert diffère peu qualitativement de l'approche ARIMA. La SCHL souhaite fournir des prévisions relativement détaillées à court terme de même qu'à plus long terme des variables relatives au logement et au marché hypothécaire. À plus long terme, l'élaboration d'un modèle économétrique d'échelle modeste constitue la première étape vers

la prévision de nombreuses variables clés. À court terme on peut faire usage des modèles ARIMA.

Tableau 6.1 Classement de quatre méthodes de prévision *

Critère	Modèles SEM	ARIMA	FT	VAR
Exigences en matière de données	4	1	2	3
Facilité de la spécification	3	1	2	4
Facilité de l'estimation	3	1	2	4
Temps/coût de la modélisation	3	1	2	4
Facilité à calculer les prévisions	4	1	2	4
Calcul de l'erreur-type des prévisions @	4	2	2	2
Ajustements subjectifs des prévisions #	1	2	3	4
Précision des prévisions à court terme	4	2	2	2
Précision des prévisions à long terme	1	3	3	3
Facilité de remise à jour	3	1	2	4
Somme des rangs	30	15	22	33

Notes :* Le classement des quatre modèles est donné selon une échelle de 1 à 4. Le meilleur modèle obtient le rang 1 et le moins bon le rang 4.

@ Il est possible, mais cela se fait rarement, d'incorporer des ajustements subjectifs des prévisions dans les modèles à séries temporelles.

On ne peut calculer les erreurs-types pour les prévisions fournies par les modèles SEM à cause de leurs ajustements subjectifs.

BIBLIOGRAPHIE

Abraham, B. (1985), « Seasonal Time Series and Transfer Function Modelling, » Journal of Business and Economic Statistics, 3, 356-361.

Aguilar, R., and B. Sandelin (1984), « Realized Supply of Owner-occupied Houses in Sweden: an Exploratory Study, » Scandinavian Housing and Planning Research, 1, 197-213.

Akaike, H. (1969), « Fitting Autoregressive Models for Prediction, » Annals of the Institute of Statistical Mathematics, 21, 243-247.

_____ (1970), « Statistical Predictor Identification, » Annals of the Institue of Mathematical Statistics, 22, 203-217.

_____ (1974), « A New Look at the Statistical Model Identification, » IEEE Transactions on Automatic Control, AC-19, 716-723.

Anderson, P.A., et J.R. Ostas (1977), « Private Credit Rationing, » New England Economic Review, 24-37.

Arnott, R. (1987), « Economic Theory and Housing, » in E.S. Mills (ed.) Handbook of Regional and Urban Economics, Vol. II: Urban Economics, Amsterdam: North-Holland.

Baer, W.C., et C.B. Williamson (1988), « The Filtering of Households and Housing Units, » Journal of Planning Literature, 3, 127-152.

Bagshaw, M. (1987), « Univariate and Multivariate ARIMA versus Vector Autoregression Forecasting, » Proceedings of American Statistical Association Business and Economics Section, 783-786.

Boehm, T.P., et K.R. Ihlanfeldt (1986), « The improvement Expenditures of Urban Homeowners: An Empirical Analysis, » American Real Estate and Urban Economics Association Journal, 14, 48-60.

Boero, G. (1990), « Comparing Ex-ante Forecasts from a SEM and VAR Model: An Application to the Italian Economy, » Journal of Forecasting, 9, 13-24.

Box, G.E.P. et G.M. Jenkins (1970), Time Series Analysis: Forecasting and Control,
San Francisco: Holden-day.

_____, et G.C. Tiao (1975), « Intervention Analysis with Applications to Economic and Environmental Problems, » Journal of the American Statistical Association, 70, 70-79.

Brady, E.A. (1973), « An Econometric Analysis of the U.S. Residential Housing Market, » in R.B. Ricks (ed.), National Housing Models, Lexington, MA: Lexington Books.

Carruthers, D.T. (1989), « Housing Market Models and the Regional Systems, » Urban Studies, 26, 214-299.

Chang, I., G.C. Tiao, et C. Chen (1988), « Estimation of Time-Series Parameters in the Presence of Outliers, » Technometrics, 30, 193-204.

Clayton Research Associates (1987), Medium and Long Term Projections of Housing Requirements in Canada, Ottawa: SCHL.

Cleveland, W.S. (1972), « The Inverse Autocorrelation of a Time Series and Their Applications, » Technometrics, 14, 277-298.

Cooper, R.L. (1972), « The Predictive Performance of Quarterly Econometric Models of the United States, » in B.G. Hickman (ed.) Econometric Models of Cyclical Behaviour, New York: Columbia University.

De Gooijer, J.G.B. Abrahams, A. Gould, et L. Robinson, « Methods for Determining the Order of Autoregressive-Moving Average Process: A Survey, » International Statistical Review, 53, 301-329.

Dent, W. (1977), « Computation of the Exact Likelihood Function of an ARIMA Process, » Journal of Statistical Computing and Simulation, 5, 193-206.

Denton, F., A. Robb, et B.G. Spencer (1986), Housing and Consumption, Ottawa: SCHL.

Dicks, M.J. (1990), « A Simple Model of Housing Starts, » Bank of England Discussion Paper No. 49.

Doan, T.A. (1990), User's Manual RATS, Evanston, IL: VAR Econometrics.

_____, R.B. Litterman, et C.A. Sims (1984), « Forecasting and Conditional Projection Using Realistic Prior Distributions, » Econometric Reviews, 3, 1-100.

Dynarski, M., et S.M. Sheffrin (1985), « Housing Purchases and Transitory Income: A Study with Panel Data », Review of Economics and Statistics, 67, 195-204.

Eckstein, O. (1981), « Econometric Models for Forecasting and Policy Analysis: the Present State of the Art, » in A.R. Sanderson (ed.) DRI Readings in Econometrics, New York: McGraw-Hill.

Emmi, P.C. (1984), « On the Duality of Opportunity and Price in a Multi-sectoral Model of an Urban Housing Market, » Environment and Planning, A, 16, 615-628.

Engle, R.F., C.W.J. Granger, et D. Kraft, « Combining Competing Forecasts of Inflation Using a Bivariate ARCH Model, » Journal of Economic Dynamics and Control, 8, 151-165.

Engle, R.F., D.M. Lilien, et M. Watson (1985), « A Dynamic Model of Housing Price Determination, » Journal of Econometrics, 28, 307-326.

Evans, M.K., Y. Haitovsky, et G.I. Treyz (1972), « An Analysis of Forecasting Properties of U.S. Econometric Models, » in B.G. Hickman, Ed. (1972), Econometric Models of Cyclical Behavior, National Bureau of Economic Research, New York: Columbia University Press.

Fackler, J.S., et S.C. Krieger (1986), « An Application of Vector Time Series Techniques to Macroeconomic Forecasting, » Journal of Business and Economic Statistics, 4, 71-80.

Fair, R.C. (1979), « An Analysis of the Accuracy of Four Macroeconomic Models, Journal of Political Economy, 87, 701-718.

_____, et R.J. Shiller (1990), « Comparing Information in Forecasts from Econometric Models, » American Economic Review, 80, 375-389.

Fisher, G.R., et M. McAleer (1981), « Alternative Procedures and Associated Tests for Significance for Non-nested Hypotheses, » Journal of Econometrics, 16, 103-119.

Foot, D.K. (1985), Housing Investment in Canadian Macroeconometric Models: An Evaluative Review, Rapport préparé pour la Division de la recherche, SCHL, Ottawa: Société canadienne d'hypothèques et de logement.

Fredland, J.E., et C.D. MacRae (1978), Econometric Models of the Housing Sector: A Policy-Oriented Survey, Washington, DC: The Urban Institute.

Friedman, B.M. (1989), « Effects of Monetary Policy on Real Economic Markets Activity » draft, Harvard University.

Fuller, W.A. (1976), Introduction to Statistical Time Series, New York: John Wiley.

Gahvari, F. (1986), « Demand and Supply of Housing in the U.S.: 1929-1978, » Economic Inquiry, 24, 333-347.

Glennon, D. (1989), « Estimating the Income, Price and Interest Elasticities of Housing Demand, » Journal of Urban Economics, 25, 219-229.

Gilderbloom, J.I. (1986), « An Analysis of Intercity Rents, » in P.L. Niebanck (ed.), The Rent Control Debate, Chapel Hill: North Carolina University Press.

Goodman, J.L., Jr., et S. Gabriel (1987), Forecasting Housing Construction: Lessons and Puzzles from Recent Years, Working Paper No. 69, for Economic Activity Section; Division of Research and Statistics, Board of Governors of the Federal Reserve System.

Goodwin, T.H. (1986), « The Impact of Credit Rationing on Housing Investment: A Multi-market Disequilibrium Approach, » International Economic Review, 27, 445-464.

Grady Economics and Associates Ltd. (1985), The State of the Art in Canadian Macroeconomic Modelling, pour la Division des prévisions économiques, Ministère des Finances.

Granger, G.W.J., et P. Newbold (1975), « Economic Forecasting: the Atheist's Viewpoint, » in G.A. Renton (ed.), Modelling the Economy, London: Heinemann Educational Books.

_____ et P. Newbold (1984), Forecasting Economic Time Series, New York: Academic Press.

_____, P. Newbold, et R. Ramanathan (1984), « Improved Methods of Combining Forecasts, » Journal of Forecasting, 3, 197-204.

Harmon, O.R. (1988), « The Income Elasticity of Demand for Single-Family Owner-Occupier Housing: An Empirical Reconciliation, » Journal of Urban Economics, 24 173-185.

Harvey, A.C. (1981), Time Series Models, Oxford: Philip Allan Publishers.

_____ (1989), Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge: Cambridge University Press.

____, et P.H.J. Todd (1983), « Forecasting Economic Time Series with Structural and Box-Jenkins Models: A Case Study with Comments, » Journal of Business and Economic Statistics, 1, 299-315.

Hatch, J.E. (1975), The Canadian Mortgage Market, Toronto: Ministry of Treasury, Economics and Intergovernmental Affairs, Queen's Printer for Ontario.

Hendershott, P.H. (1988), « Home Ownership and Real House Prices: Sources of Change, 1965-1985 », Housing Finance Review, 7, 1-18.

_____ et M.T. Smith (1988), « Housing Inventory Change and the Adjustment Role of Existing Structures, 1961-1985, » AREUEA Journal, 16, 364-378.

Hillmer, S. (1984), « Monitoring and Adjusting Forecasts in the Presence of Additive Outliers, » Journal of Forecasting, 3, 205-215.

Hillmer, S.C., et G.C. Tiao (1979), « Likelihood Function of Stationary Multiple Autoregressive Moving Average Models, » Journal of American Statistical Association, 74, 652-660.

Howrey, E.P., L.R. Klein, et M.D. McCarthy (1974), « Notes on Testing the Predictive Performance of Econometric Models, » International Economic Review, 15, 366-383.

Hsiao, C. (1979), « Autoregressive Modelling of Canadian Money and Income, » Journal of American Statistical Association, 74, 553-560.

_____ (1981), « Autoregressive Modelling and Money-Income Causality Detection, » Journal of Monetary Economics, 7, 85-106.

Jaffee, D.M., et K.T. Rosen (1978), « Estimates of Effectiveness of Stabilization Policies for the Mortgage and Housing Markets, » Journal of Finance, 33, 933-946.

Jaffee, D., et K. Rosen (1979), « Mortgage Credit Availability and Residential Construction, » Brookings Papers on Economic Activity, No.2, 333-366.

Jud, D.J., et J. Frew (1986), « Real Estate Brokers, Housing Prices, and the Demand for Housing, » Urban Studies, 23, 21-31.

Kahn, G.A. (1989), « The Changing Interest Sensitivity of the U.S. Economy, » Economic Review, 13-34.

Klein, L.R. (1950), Economic Fluctuations in the United States, 1921-1941, New York: John Wiley.

_____ (1968), An Essay on the Theory of Economic Prediction, Helsinki: Yrjö Jahnssonin Säätiö.

_____ (1971), « Forecasting and Policy Evaluation Using Large Scale Econometric Models: the State of the Art. » in M.D. Intriligator (ed.), Frontiers of Quantitative Economics, Amsterdam: North Holland.

_____ (1982), « Comments and Discussion, » following Christopher A. Sims, « Policy Analysis with Econometric Models, » Brookings Papers on Economic Activity, No.1, 164.

Krumm, R. (1987), « Intertemporal Tenure Choices, » Journal of Urban Economics, 22, 263-275.

Litterman, R.B. (1979), « Techniques of Forecasting Using Vector Autoregressions, » Workign Paper 115, Federal Reserve Bank of Minneapolis.

_____ (1980), « A Bayesian Procedure for Forecasting with Vector Autoregressions, » Working Paper, MIT.

_____ (1984), « Specifying Vector Autoregressions for Macroeconomic Forecasting, » Staff Report 92. Federal Reserve Bank of Minneapolis.

_____ (1986), « Forecasting with Bayesian Vector Autoregression - Five Years of Experience, » Journal of Business and Economic Statistics, 4, 25-38.

Liu, L.M., et D.M. Hannssens (1982), « Identification of Multiple Input Transfer Function Models, » Communications in Statistics, 11, 297-314.

_____ et al. (1986). The SCA Statistical System: Reference Manual for Forecasting and Time Series Analysis, DeKalb, IL: Scientific Computing Associates.

Lütkepohl, H. (1985), « Comparison of Criteria for Estimating the Order of a Vector Autoregressive Process, » Journal of Time Series Analysis, 6, 35-52.

Makridakis, S. et al. (1982), « The Accuracy of Extrapolation (Time Series Methods: Results of a Forecasting Competition, » Journal of Forecasting, 1, 111-153.

Malpezzi, S., L. Ozanne, et T.G. Thibodeau (1987), « Microeconomic Estimates of Housing Depreciation, » Land Economics, 63, 372-385.

Manning, C.A. (1986), « Intercity Differences in Home Price Appreciation, » Journal of Real Estate Research, 1, 45-66.

McAleer, M., et M.H. Pesaran (1986), « Statistical Inference in Non-nested Econometric Models, » Applied Mathematics and Computation, 271-311.

McNees, S.K. (1986), « Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts, »

Journal of Business and Economic Statistics, 4, 5-23 (avec commentaires).

_____ (1988), « How Accurate Are Macroeconomic Forecasts?, » New England Economic Review, Juillet-août 15-36.

Mills, T.C., et M.J. Stephenson (1987), « A Time Series Forecasting System for the UK Money Supply, » Economic Modelling, 4, 355-369.

Mutchler, J.E., et L.J. Krivo (1989), « Availability and Affordability: Household Adaptation to a Housing Squeeze, » Social Forces, 68, 241-261.

Muth, Richard (1986), « The Supply of Mortgage Lending, » Journal of Urban Economics, 19, 88-106.

Muth, R.F. (1988), « Housing Market Dynamics, » Regional Science and Urban Dynamics Economics, 18, 345-356.

Nelson, C.R. (1972), « The Prediction Performance of the FRB-MIT - Penn Model of the U.S. Economy, » American Economic Review, 92, 902-917.

_____ (1984), « A Benchmark for the Accuracy of Econometric Forecasts of GNP, » Business Economics, 19, 52-58.

Newbold, P. (1974), « The Exact Likelihood Function for a Mixed Autoregressive Moving, Average Process, » Biometrika, 61, 423-426.

_____ (1983a), « ARIMA Model Building and the Time Series Analysis Approach to Forecasting, » Journal of Forecasting, 2, 23-35.

_____ (1983b), « Model Checking in Time Series Analysis, » in A. Zellner (ed.), Applied Time Series Analysis, Washington, D.C.: U.S. Bureau of the Census (avec commentaires de Ledolter and Pierce).

Olsen, E. (1987), « Demand and Supply of Housing Service: A Critical Survey of the Empirical Literature, » in E.S. Mills (ed.), Handbook of Regional and Urban Economics, Vol. II: Urban Economics, Amsterdam: North-Holland.

Park, S.-B. (1989) Applied Time Series Analysis for Business and Economics, Mimeo., Carleton University.

Pankratz, A., (1983), Forecasting with Univariate Box-Jenkins Models, New York: John Wiley.

_____ (1989), « Time Series Forecasts and Extra-Model Information, » Journal of Forecasting, 8, 75-83.

Pelser, R.B., et L.B. Smith (1985), « Homeownership Returns, Tenure Choice and Inflation, » AREUEA Journal, 13, 343-360.

Pesaran, M.H. (1974), « On the general problem of model selection, » Review of Economic Studies, 43, 153-171.

Poskitt, D.S. (1989), « A Method for the Estimation and Identification of Transfer Function Models, » Journal of the Royal Statistical Society, B, 51, 29-46.

_____, et A.R. Tremayne (1981), « An Approach to Testing Linear Time Series Models, » Annals of Statistics, 5, 974-986.

Puri, A., et J. Van Lierop (1988), « Forecasting Housing Starts, » International Journal of Forecasting, 4, 125-134.

Quigley, J.M. (1987), « Interest Rate Variations, Mortgage Prepayments and Household Mobility, » Review of Economics and Statistics, 69, 636-643.

Quinn, B.G. (1988), A Note on AIC Order Determination for Multivariate Autoregressions, » Journal of Time Series Analysis, 9, 241-245.

Raynauld, J., and J. Simonato (1988), « Seasonal BVAR Models, » Proceedings of American Statistical Association Business and Economic Statistics Section, 319-324.

_____ (1990), « Seasonal BVAR Models: A Search Along Some Time Domain Priors, » Paper presented at the CRDE/Journal of Econometrics Conference on Seasonality and Econometric Models.

Ricks, R.B. (1973), National Housing Models: Application of Econometric Techniques to Problems of Housing Research, Lexington, MA: Lexington Books.

Rosen, K.T., et L.B. Smith (1986), « The Resale Housing Market », AREUEA Journal, 14, 510-524.

Rosen, S., et R. Topel (1986), A Time-Series Model of Housing Investment in the U.S., Working Paper No. 1818, New York: National Bureau of Economic Research.

Rosenthal, L. (1986), « Regional House Price Interactions in the U.K., 1971-1981: A Cross-Spectral Analysis, » Applied Economics, 18, 1011-1023.

Runkle, D.E. (1987), « Vector Autoregressions and Reality, » Journal of Business and Economic Statistics, 5, 437-454 (avec commentaires).

Shibata, R. (1985), « Various Model Selection Techniques in Time Series Analysis, » in E.J. Hannan, P.R. Krishnaiah, et M.M. Rao (eds.), Handbook of Statistics, Vol. 5, New York: Elsevier Science Publishers, 179-187.

Sims, C.A. (1980), « Macroeconomics and Reality, » Econometrica, 48, 1-48.

_____ (1982), « Policy Analysis with Economic Models, » Brookings Papers on Economic Activity, No. 1.

Sklarz, M.A., N.G. Miller, et W. Gersch (1987), « Forecasting Using Long-Order Autoregressive Processes: An Example Using Housing Starts, » AREUEA Journal, 15, 374-388.

Smith L.B. (1974), The Postwar Canadian Housing and Residential Mortgage Markets and the Role of Government, Toronto, University of Toronto Press.

_____, L.B., K.T. Rosen, et G. Fallis (1988), « Recent Developments in Economic Models of Housing Markets, » Journal of Economic Literature, 26, 29-64.

_____, et G.R. Sparks (1970), « The Interest Sensitivity of Canadian Mortgage Flows, » Canadian Journal of Economics, 3, 407-421.

Stansell, S.R., et A.C. Mitchell (1985), « The Impact of Credit Rationing on the Real Sector: A Study of the Effect of Mortgage Rates and Terms on Housing Starts, » Applied Economics, 17, 781-800.

Stekler, H.O. (1968), « Forecasting with Econometric Models: An Evaluation, » Econometrica, 36, 427-463.

Stilner, K.R., et D.R. Barton (1990), « Econometric Models and Construction Forecasting, » U.S. Department of Commerce Construction Review, 10-20.

Stokes, E. (1987), Canadian Economic and Fiscal Model (CEFM): Basic Structure and Specification, Division des analyses et des prévisions économiques, Ministère des Finances.

Thom, R. (1985), « The Relationship between Housing Starts and Mortgage
Availability, » Review of Economics and Statistics, 67, 693-696.

Todd, R.M. (1984), « Improving Economic Forecasting with Bayesian Vector Autoregression, » Federal Reserve Bank of Minneapolis Quarterly Review, 18-29.

Topel, R., et S. Rosen (1988), « Housing Investment in the United States, » Journal of Political Economy, 96, 718-740.

Tsay, R.S., et G.C. Tiao (1984), « Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models, » Journal of American Statistical Association, 79, 84-96.

Turner, M.A., et R.J. Struyk (1985), Urban Housing in the 1980s: Markets and Policies, Washington, DC: Urban Institute Press.

Vandaele, W. (1983), Applied Time Series and Box-Jenkins Models, New York: Academic Press.

Wang, M., et H. Ma (1981), « Multiple Time Series Models for Housing and the Mortgage Rate, » Proceedings of the American Statistical Association Business and Economic Statistics Section, 418-423.

Weicher, J.C., et T.G. Thibodeau, T.G. (1988), « Filtering and Housing Markets: An Empirical Analysis, » Journal of Urban Economics, 23, 21-40.

Wiliams, R.A. (1984), « An Australian Housing Model, » Economic Record, 60, 143-155.

Zellner, A., et F. Palm (1979), « Time Series Analysis and Simultaneous Equation Econometric Models, » Journal of Econometrics, 2, 17-54.

Zumpano, L.V., P. Rudolph, et D.C. Cheng (1982), « The Demand and Supply of Mortgage Funds and Mortgage Loans Terms, » AREUEA Journal, 14.